# MEASUREMENT OF HIGGS BOSON PRODUCTION IN ASSOCIATION WITH A TOP QUARK-ANTIQUARK PAIR IN FINAL STATES WITH LEPTONS AND HADRONIC TAUS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Zhengcheng Tao

August 2019

MEASUREMENT OF HIGGS BOSON PRODUCTION IN ASSOCIATION WITH A TOP QUARK-ANTIQUARK PAIR IN FINAL STATES WITH LEPTONS AND HADRONIC TAUS

Zhengcheng Tao, Ph.D.

Cornell University 2019

A measurement of the cross section of the Higgs boson production associated with a pair of top quarks ($t\bar{t}H$) in proton-proton collisions at a center-of-mass energy of 13 TeV is presented. The measurement targets $t\bar{t}H$ final states with electrons, muons, and hadronically decaying tau leptons. Data used in this measurement are recorded by the Compact Muon Solenoid experiment at the CERN Large Hadron Collider in 2016 and 2017, corresponding to an integrated luminosity of 35.9 $fb^{-1}$ and 41.5 $fb^{-1}$, respectively. The measured signal rate is $0.96^{+0.34}_{-0.31}$ times the expected production rate in the standard model. The observed (expected) significance of the $t\bar{t}H$ signal is $3.2\sigma$ ($4.0\sigma$).

A proposed FPGA-based track finding system for the CMS Level-1 trigger upgrade in preparation for the High-Luminosity LHC is also presented in this thesis.

## BIOGRAPHICAL SKETCH

Zhengcheng Tao was born in Changzhou, China. He spent two years as an undergraduate student at Sichuan University in Chengdu, China, before he transferred to the University of Wisconsin-Madison, where he received his Bachelor of Science in physics in May 2013. At UW-Madison, Zhengcheng worked with Doctor Christopher Wendt on the IceCube Neutrino Observatory and worked with Professor Sau Lan Wu on the ATLAS experiment at the CERN Large Hadron Collider (LHC).

Zhengcheng started as a graduate student in physics at Cornell University in August 2013. He has been working on the CMS experiment at the LHC with Professor Anders Ryd and Professor Peter Wittich. Most of his work on the physics analysis and detector upgrades are presented in this thesis. Zhengcheng is based in Ithaca, NY during his graduate studies, with occasional travels to CERN to attend meetings and participate in the CMS strip tracker operation, and with frequent travels to the Toronto area to visit his then-girlfriend-now-wife.

Zhengcheng will continue his research at the energy frontier as a postdoctoral associate with the ATLAS group at the University of British Columbia in Vancouver, Canada.

**ACKNOWLEDGEMENTS**

First of all, I would like to thank my advisor, Anders Ryd, for his guidance and support over the years. It has truly been a pleasure and privilege to have Anders as my advisor. I have learned a lot from Anders about collider phenomenology, particle detectors, programming, joking, and much more. His encouragement and passion for particle physics motivate me to continue the adventure as a postdoctoral researcher at the energy frontier.

I would also like to thank Peter Wittich. I learned the standard model of particle physics from Peter and have had a lot of fun working with Peter on the Tracklet project. I want to thank Peter for his insights and ideas on the project, and especially for driving the HLS firmware effort. I am also very grateful for his advice and support, and for practically being my second advisor. Additionally, I would like to thank other faculty members of the Cornell CMS group, Jim Alexander, Julia Thom, and Ritchie Patterson, for fostering a supportive research environment.

One of the greatest bonuses of working at Cornell is being able to benefit from the theory group upstairs. I want to thank in particular Maxim Perelstein for serving on my committee, for his collider physics seminar, and for running the espresso club, which literally fueled my graduate studies.

The $t\bar{t}H$ analysis would not be possible without the excellent performance of the CMS detector and the support of the physics object groups. It is a result of dedicated and brilliant work of many colleagues. I would like to thank everyone in the $t\bar{t}H$ multi-lepton and tau working group, in particular Aurelijus Rinkevicius for many useful discussions, Christian Veelken for his expertise in tau reconstruction and identification and for leading the analysis in tau categories, Marco Peruzzi and Giovanni Petrucciani for the lepton MVA and fake

factor measurements. Thomas Strebler and Matthias Wolf contributed significantly to the analysis in tau categories with 2016 data. Alexandra Carvalho was responsible for training the event-level BDTs for the 2017 analysis and helped greatly on the statistical analysis. I also want to thank Karl Ehataht, Cristina Perez, and Binghuan Li regarding synchronizations of the object and event selection.

I enjoyed very much working on the Tracklet project, not only because the project itself is very exciting but also because of the awesome people I worked with. In addition to Anders and Peter, I would like to thank especially Jorge Chaves, Margaret Zientek, and Charlie Strohman for firmware and hardware development, Louise Skinnari for software simulation and performance studies, Derek Cranshaw for taking everything over after Margaret and I are gone. I would also like to thank all my collaborators from Northwestern University, the Ohio State University, Rutgers University, Notre Dame University, and University of Colorado Boulder.

I would like to thank my friends and fellow graduate students in the CMS group as well. I could not have done this alone. Thanks particularly to Abhisek Datta, Andre Frankenthal, Dan Quach, Derek Cranshaw, Divya Gadkari, Jennifer Chu, Jordan Tucker, Jorge Chaves, Kevin McDermott, Margaret Zientek, Nathan Mirman, Shao Min Tan, Shaun Hogan, Susan Dittmer, Tres Reid.

Finally, I would like to thank my parents, Guihua Zhou and Xihua Tao, and my wife, Jing Min, for their love and support for my educational and academic endeavors.

## TABLE OF CONTENTS


Biographical Sketch . . . iii
Acknowledgements . . . iv
Table of Contents . . . vi
List of Tables . . . ix
List of Figures . . . xi

**1 Introduction** **1**
1.1 Higgs boson at the LHC . . . 1
1.2 Associated Higgs boson production with a top quark pair . . . 3

**2 Standard model and Higgs couplings** **7**
2.1 Boson and fermion fields . . . 7
2.2 The Higgs mechanism . . . 10
2.3 Yukawa couplings . . . 13
2.4 Measure the top-Higgs coupling . . . 15
2.4.1 Parameterization . . . 16
2.4.2 Indirect measurements . . . 17
2.4.3 Direct measurements . . . 19

**3 Experimental apparatus** **21**
3.1 Large Hadron Collider . . . 21
3.2 Compact Muon Solenoid detector . . . 23
3.2.1 Coordinate system . . . 25
3.2.2 Solenoid . . . 25
3.2.3 Silicon tracker . . . 26
3.2.4 Electromagnetic calorimeter . . . 30
3.2.5 Hadron calorimeter . . . 33
3.2.6 Muon system . . . 36
3.2.7 Trigger and data acquisition system . . . 39

**4 Object reconstruction** **44**
4.1 Particle-flow reconstruction . . . 44
4.2 Tracks . . . 46
4.3 Vertices and beam spot . . . 50
4.4 Muons . . . 50
4.5 Electrons . . . 53
4.6 Jets . . . 57
4.7 b tagging . . . 58
4.8 Hadronic taus . . . 59
4.9 Missing transverse momentum . . . 61

**5 Measurement of $t\bar{t}H$ production cross section in final states with leptons and hadronic taus** **62**
5.1 Data sets and Monte Carlo samples . . . 64
5.2 Object selection . . . 67
5.2.1 Muons and electrons . . . 67
5.2.2 Hadronic taus . . . 73
5.2.3 Jets and missing transverse momentum . . . 74
5.3 Event selection . . . 76
5.3.1 Triggers . . . 77
5.3.2 Lepton and $\tau_h$ multiplicity . . . 78
5.3.3 Kinematics . . . 81
5.3.4 Charge requirements . . . 81
5.3.5 Jet and b tagging requirements . . . 82
5.3.6 Missing transverse momentum . . . 83
5.3.7 Lepton mass veto . . . 84
5.4 Modeling of signal and background processes . . . 86
5.4.1 Background from misidentified leptons and $\tau_h$'s . . . 87
5.4.2 Background from lepton charge misidentification . . . 91
5.4.3 Background from photon conversion . . . 92
5.4.4 Irreducible backgrounds . . . 93
5.4.5 Data-to-MC corrections . . . 94
5.4.6 Background control regions . . . 99
5.5 Signal extraction . . . 103
5.5.1 Discriminating observables . . . 103
5.5.2 Statistical analysis . . . 107
5.6 Systematic uncertainties . . . 113
5.7 Results . . . 118

**6 CMS Level-1 track trigger upgrade for the HL-LHC** **124**
6.1 High-Luminosity LHC . . . 124
6.2 CMS Phase-2 tracker . . . 127
6.3 Tracklet approach for L1 tracking . . . 130
6.3.1 Algorithm . . . 131
6.3.2 Implementation . . . 138
6.3.3 Hardware demonstrator . . . 143
6.3.4 L1 tracking performance . . . 148
6.4 Summary . . . 151

**7 Conclusion** **153**

**A Auxiliary measurements for reducible background estimations** **155**
A.1 Lepton fake factor . . . 155
A.1.1 Fakeable lepton definition tuning . . . 155
A.1.2 Measurement of lepton fake factor . . . 158

A.2 Hadronic tau fake factor . . . . . 162
A.3 Lepton charge misidentification rate . . . . . 166

**B Potential improvements in event categories with hadronic taus 168**
B.1 Tau decays as spin analyzers . . . . . 168
B.2 Spin correlations in signal and background processes . . . . . 171
B.3 Application and challenges . . . . . 172

**Bibliography 175**

## LIST OF TABLES


2.1 Representations of SM gauge group for fermions. . . . . 10

5.1 List of data sets in 2017 used in this analysis. . . . . 65
5.2 Effective areas $\mathcal{A}$ for muons (left) and electrons (right). . . . . 68
5.3 Loose, fakeable, and tight selection criteria for muons. A hyphen (−) indicates the selection is not applied. The relaxed selection criteria for the fakeable selection in case of prompt-$\mu$ BDT > 0.90 are placed in parentheses. . . . . 71
5.4 Loose, fakeable, and tight selection criteria for electrons. A hyphen (−) indicates the selection is not applied. The relaxed selection criteria for the fakeable selection in case of prompt-$\mu$ BDT > 0.90 are placed in parentheses. The requirements on $\sigma_{i\eta i\eta}$ are different in the barrel and endcap regions. The numbers in the curly brackets are separated by | and refer to $|\eta| < 1.479$ and $|\eta| \geq 1.479$, respectively. . . . . 72
5.5 Definition of the electron MVA ID loose working point. . . . . 72
5.6 Fakeable, and tight selection criteria for hadronic taus. . . . . 74
5.7 High-level triggers used to record events selected in the $1\ell + 2\tau_{\mathrm{h}}$, $2\ell$ss, $2\ell$ss + $1\tau_{\mathrm{h}}$, $2\ell + 2\tau_{\mathrm{h}}$, $3\ell$, $3\ell + 1\tau_{\mathrm{h}}$, and $4\ell$ channels. . . . . 77
5.8 Event selections applied in the $2\ell$ss, $2\ell$ss$+1\tau_{\mathrm{h}}$, $1\ell+2\tau_{\mathrm{h}}$ and $2\ell+2\tau_{\mathrm{h}}$ categories. . . . . 79
5.9 Event selections applied in the $3\ell$, $3\ell + 1\tau_{\mathrm{h}}$ and $4\ell$ categories. . . 80
5.10 $p_{\mathrm{T}}^{\mathrm{miss}}$ filters applied to collision data and to simulated events. A hyphen (−) indicates that the filter is not applied. . . . . 84
5.11 Trigger efficiency scale factors applied to simulated events for $2\ell$ss$(+1\tau_{\mathrm{h}})$, $3\ell(+1\tau_{\mathrm{h}})$, and $4\ell$ categories. . . . . 96
5.12 Scale factors of $\tau_{\mathrm{h}}$ selection efficiencies for Very-loose, Loose, and Medium $\tau_{\mathrm{h}}$ MVA ID working point. . . . . 99
5.13 Variables used in the event-level BDT classifiers in each event category to discriminant the $\mathrm{t\bar{t}H}$ signal against $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$, and $\mathrm{t\bar{t}}$+jets background processes. . . . . 105
5.14 Summary of the main sources of systematic uncertainty and their impact on the measured signal rate. . . . . 117
5.15 Event yields in all categories after the maximum likelihood fit. The "SM expectation" in the table includes contributions from the $\mathrm{t\bar{t}H}$ process. . . . . 122
5.16 Measure and expected $\mathrm{t\bar{t}H}$ signal strength modifiers in each event category and with all seven categories combined using data recorded in 2017. The 2017 result is also combined with that obtained using data recorded in 2016. The result combining two years is shown in the last row of the table. . . . . 123

6.1 FPGA resource usages for a full-sector project report by Vivado (top row). Fractions of these resources over what are available on different FPGAs, including the Virtex-7 used for the demonstrator, are shown in the following rows. . . . . 145
6.2 The demonstrator latency model. Estimated latencies for each processing step and for the total system are shown. . . . . 147

A.1 Triggers used to record events for the measurement of the lepton misidentification rate. . . . . 156
A.2 Fractions of the same-sign events in each category expressed as sums of the corresponding charge misidentification rates. . . . . 167
A.3 Electron charge misidentification rates in $p_{\mathrm{T}}$ and $\eta$ bins. . . . . 167

B.1 Main $\tau^-$ decay modes and branching ratios. Branching ratio values are taken from Ref. [62]. Decay modes of $\tau^+$ are the charge conjugate of the $\tau^-$ modes. . . . . 169

## LIST OF FIGURES


1.1 Main Higgs boson production modes at the LHC. From left to right: gluon-gluon fusion (ggF), vector boson fusion (VBF), associated Higgs production with vector boson (VH), associated Higgs production with a pair of top quarks ($t\bar{t}H$). . . . . 1

1.2 The SM Higgs boson production cross sections as a function of the center-of-mass energy for proton-proton collisions [3]. The production mode VBF is indicated as qqH in the plot. In addition to the four main Higgs production modes, associated Higgs productions with a bottom quark pair (bbH) and with a single top quark (tH) are also shown. . . . . 2

1.3 Branching ratios of Higgs boson decay modes around $m_H$ = 125 GeV [4]. . . . . 3

1.4 Main Feynman diagrams for $t\bar{t}H$ productions at the LHC. . . . . 4

1.5 A simulated event for $t\bar{t}H$ in multi-lepton final states [24]. . . . . 6

2.1 Left: best fit values of Higgs coupling modifiers in the 6-$\kappa$-parameter framework for ATLAS (blue) and CMS (red) individually and for the combination of data from the two experiments (black). It is assumed that no BSM particles contributed to the loop processes. Right: best fit values of Yukawa coupling strengths as a function of particle masses based the combined results of ATLAS and CMS. The SM prediction is indicated with the blue dashed line. Data are also fitted with a linear function as indicated by the red solid line with 68% and 95% confidence level (CL) bands. Plots are taken from Ref. [6]. . . . . 18

2.2 Fit results of Yukawa coupling modifiers with $\kappa_t$ directly constrained by the $t\bar{t}H$ process under two parameterizations: one allows non-zero BSM Higgs branching ratio ($|\kappa_V| \leq 1$ and $B_{BSM} \geq 0$); the other does not ($B_{BSM} = 0$) [6]. Results based on ATLAS (blue) and CMS (red) data separately, as well as on the combined data (black), are shown in the plots. . . . . 19

3.1 The accelerator complex at CERN including the LHC [29]. The diagram also shows the four major experiments on the LHC. . . 22

3.2 The overall layout of CMS [32]. Detector components from innermost to outermost: silicon tracker, electromagnetic calorimeter, hadron calorimeter, superconducting solenoid, and muon detector. . . . . 24

3.3 Schematic cross section through the CMS tracker in the $r$-$z$ plane [32]. Each sensor module is represented by a line segment. 26

3.4 Layout of the CMS electromagnetic calorimeter [32]. . . . . 30

3.5 Schematic of a quarter of the CMS detector in $r$-$z$ view [32]. The HCAL subsystems are labeled in the diagram. . . . . 33

3.6 Cross section of a quarter of the CMS muon system in $r$-$z$ view [40]. 37
3.7 Left: the schematic of a DT chamber showing orthogonal orientations of the drift cells in different layers. Right: a DT cell with the drift lines [40]. . . . . 37
3.8 Left: cut-away view of a CSC with 7 trapezoidal cathode strip panels and 6 gas gaps with anode wires [32]. Right: schematic of a RPC in barrel [40]. . . . . 39
3.9 Data-flow for the Level 1 Trigger [41]. . . . . 41

4.1 A transverse view of a slice of the CMS detector showing the interactions of different types of particles identified by the particle-flow algorithm with subdetectors [44]. . . . . 45
4.2 Left: resolution of the relative $p_T$ as a function of simulated track $p_T$. Right: resolution of the transverse impact parameter $d_0$ as a function of simulated track $\eta$. Both resolution plots are based on simulated $t\bar{t}$ events with an average pileup of 35. Performances in both 2016 (blue) and 2017 (red) detector conditions are shown [46]. . . . . 47
4.3 Muon medium ID efficiency as a function of muon $p_T$ (left) and $\eta$ (right) for 2017 data (black) and simulation (blue) [53]. The denominator for the efficiency calculation is all tracker tracks with $p_T > 20$ GeV. . . . . 52
4.4 Receiver operating characteristic curves of the electron BDT-based identification [56]. . . . . 56

5.1 Closure tests using simulated $t\bar{t}$ + jets events in the $1\ell + 2\tau_h$ category. Distributions of the number of jets (left) and leading $\tau_h$ candidate $p_T$ (right) extrapolated from the AR using the FF method ("TT(FR)", in blue) are compared to those obtained directly from the SR ("Fakes_MC", in black). . . . . 91
5.2 Lepton selection efficiencies as functions of $p_T$ of loose muon and electron to pass the corresponding tight selection criteria. Muon efficiencies in $0 \leq |\eta| < 0.4$ (left) and $0.4 \leq |\eta| < 0.8$ (right) regions are shown on the top row. Electron efficiencies in $0 \leq< |\eta| < 0.74$ and $0.74 \leq |\eta| < 1.479$ regions are on the bottom row. Efficiencies measured in recorded data and MC simulated Drell-Yan (DY) samples are shown in black and blue markers, respectively. Data-to-MC ratios are also shown. . . . . 98
5.3 Event yields (left) and the distribution of di-$\tau_h$ invariant mass (right) in the $1\ell + 2\tau_h$ss control region. . . . . 100
5.4 The $E_T^{miss}$ distribution (left) and event yields in $2\ell$ss subcategories (right) in the $t\bar{t}W$ control region. . . . . 101
5.5 Distributions of $E_T^{miss}$ (left) and dilepton mass that is closest to the mass of a Z boson in the $t\bar{t}Z$ control region. . . . . 102

5.6 Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and the number of selected jets (right) in the WZ control region. . . . . 102
5.7 Distributions of discriminating observables after the fit in $2\ell$ss (left) and $3\ell$ (right) categories. . . . . 120
5.8 Distributions of discriminating observables after the fit in $\mathrm{t\bar{t}W}$ (left) and $\mathrm{t\bar{t}Z}$ (right) control regions. . . . . 120
5.9 Post-fit distributions of discriminating observables in $1\ell + 2\tau_{\mathrm{h}}$ (top left), $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$ (top right), $3\ell + 1\tau_{\mathrm{h}}$ (bottom left), and $2\ell + 2\tau_{\mathrm{h}}$ (bottom right) categories. . . . . 121
5.10 Measured $\mathrm{t\bar{t}H}$ signal strength modifiers for individual event categories and for the combination of all event categories using 41.5 $\mathrm{fb}^{-1}$ data recorded in 2017. . . . . 123

6.1 LHC/HL-LHC operation and upgrade timeline [96]. The center-of-mass energy and instantaneous luminosity during each period are indicated in the figure. The nominal instantaneous luminosity of the LHC is $1.0 \times 10^{34}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The total amount of data in terms of integrated luminosities at the end of each operation period are also shown. . . . . 125
6.2 Left: L1 single muon trigger rate as a function of trigger $p_{\mathrm{T}}$ thresholds; Right: the efficiency of the L1 single muon trigger with a threshold $p_{\mathrm{T}} > 20$ GeV as a function of simulated muon $p_{\mathrm{T}}$ [95]. The L1 standalone muon trigger (red) and the trigger combining L1 muon and L1 track information (black) are shown in the pseudo-rapidity range of $|\eta| < 1.1$ (solid circles) and $1.1 \leq |\eta| \leq 2.4$ (hollow circles). . . . . 126
6.3 Top: the concept of $p_{\mathrm{T}}$ modules. Hits in the two closely spaced sensors are correlated to provide on-detector $p_{\mathrm{T}}$ discrimination. Bottom: strip-strip modules (left) and pixel-strip modules (right). All plots are taken from Ref. [98]. . . . . 128
6.4 The layout of the proposed CMS Phase-2 tracker in the *r*-*z* view [98]. One quarter of the tracker is shown in the plot. The blue and red modules are pixel-strip and strip-strip modules for the silicon strip tracker, respectively. The inner pixel tracker modules are indicated in orange and green, and they are not used in the L1 track finding. . . . . 129
6.5 An overview of the Tracklet algorithm. Left: the seeding step, in which a tracklet is formed from a pair of stubs (red) in adjacent layers/disks. Middle: the projection step, in which projected positions to other layers/disks are estimated based on the seed. Right: the matching and fitting step, in which stubs in other layers/disks are compared to the projected trajectory. The matched stubs (green) and the seeding stubs (red) are grouped together to compute the track parameters by a $\chi^2$ fit. . . . . 131

6.6 The definition of processing sectors in $x$-$y$ view of the tracker. The green highlighted area is one of the nine $\phi$ sectors. Dotted circle in gray indicated the critical radius $r_c$. . . . . 132
6.7 Tracking efficiency in single muon events as a function of $\eta$ for different seeding pairs. . . . . 133
6.8 An illustration of virtual modules in a pair of seeding layers. A track with its $p_T$ below the threshold (red) would bend more than a track with a larger $p_T$ (green) in the CMS magnetic field. Therefore only some combinations of inner and outer VMs are worth spending computing resources to process. . . . . 134
6.9 Distributions of number of tracks per single muon event before (black) and after (red) the duplicate removal. The plot is based on simulated events, each of which contains one muon with $p_T$ between 2 and 10 GeV. . . . . 137
6.10 Resolutions of relative $p_T$, $\phi_0$, $\eta$, and $z_0$ of L1 tracks obtained using full precision floating point calculation (solid red) and using the integer emulation (dashed blue). These plots are produced based on simulated tracks satisfying $2 < p_T < 10$ GeV and $|\eta| < 2.4$ in single muon events without pileup. . . . . 139
6.11 An overview of processing steps (red) and memory types (blue) of the project. Grey arrow indicate the data dependencies among the steps. . . . . 140
6.12 Left: a schematic overview of the Tracklet demonstrator. Three CTP7 boards are used as sector processing boards, and one CTP7 board is used as both the input stub source and as the output track sink. Estimated data transmission rates on the inter-board links are indicated as well. Right: a photo of the demonstrator system in a $\mu$CTA crate. . . . . 144
6.13 Tracker regions covered by the barrel-only (solid blue) and hybrid-region (dashed green) projects for the demonstrator. The flat tracker geometry shown here is the layout assumed by the version of the Tracklet firmware that was built for the demonstrator. . . . . 146
6.14 Comparisons of final track parameters $p_T$ (left) and $z_0$ (right) calculated in emulation (black) and in firmware simulation (red) in simulated $t\bar{t}$ events with an average pileup of 200. The barrel-only project plots are shown in the top row, and plots of the hybrid-region project are shown in the bottom. . . . . 148
6.15 Tracking efficiency in simulated $t\bar{t}$ events with an average pileup of 0 (black), 140 (red), and 200 (blue). Efficiencies as a function of track $p_T$ and $\eta$ are shown in the top and bottom row, respectively. Plots in the left column are for low $p_T$ range ($2 < p_T < 8$ GeV), and those on the right are for high $p_T$ range ($8 < p_T < 100$ GeV). . 149

6.16 Resolutions of final track parameters $p_T$ (top left), $\phi_0$ (top right), $\eta$ (bottom left), and $z_0$ (bottom right) as a function of track $|\eta|$. Simulated $t\bar{t}$ events with an average of 200 pileup are used for the plots. Low $p_T$ range ($2 < p_T < 8$ GeV) and high $p_T$ range ($8 < p_T < 100$ GeV) are indicated by blue and black markers, respectively. Resolutions shown in these plots correspond to intervals that encompass 68% (solid marker) and 90% (open marker) of tracks in the corresponding track parameter distribution. . . . 150

A.1 Distributions of $m_T^{fix}$ in the pass (left) and fail (right) regions for muons satisfying $15 <$ cone-$p_T < 20$ GeV and $|\eta| > 1.2$. . . . . 159
A.2 Distributions of $m_T^{fix}$ in the pass (left) and fail (right) regions for electrons satisfying $25 <$ cone-$p_T < 35$ GeV and $|\eta| < 1.479$. . . . . 160
A.3 Measured muon fake factors as functions of $p_T$ in the barrel (left) and endcap (right) regions. . . . . 161
A.4 Measured $\tau_h$ fake factors as functions of $p_T$ in the barrel (left) and endcap (right) regions. . . . . 161
A.5 Measured $\tau_h$ fake factors as functions of $p_T$ in the barrel (left) and endcap (right) regions. The tight $\tau_h$ definition for $2\ell$ss + $1\tau_h$ and $3\ell + 1\tau_h$ categories with the Loose working point of $\tau_h$ MVA ID is used. . . . . 163
A.6 Measured $\tau_h$ fake factors as functions of $p_T$ in the barrel (left) and endcap (right) regions. The tight $\tau_h$ definition for $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ categories with the Medium working point of $\tau_h$ MVA ID is used. . . . . 163
A.7 Ratio of $\tau_h$ fake factors measured in data to those in MC simulation in the barrel (left) and endcap (right) regions. The ratio is fitted to a linear function $c_0 + c_1 p_T$. The tight $\tau_h$ definition for $2\ell$ss + $1\tau_h$ and $3\ell + 1\tau_h$ categories with the Loose working point of $\tau_h$ MVA ID is used. . . . . 165
A.8 Ratio of $\tau_h$ fake factors measured in data to those in MC simulation in the barrel (left) and endcap (right) regions. The ratio is fitted to a linear function $c_0 + c_1 p_T$. The tight $\tau_h$ definition for $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ categories with the Medium working point of $\tau_h$ MVA ID is used. . . . . 165

B.1 Distributions of $\Upsilon$ at generator level for $\tau_L^-$ and $\tau_R^+$ (left) and for $\tau_R^-$ and $\tau_L^+$ (right). Plots are taken from Ref. [105]. . . . . 170
B.2 Distributions of generator level $\tau^+$ and $\tau^-$ visible energies in the mother boson rest frame, normalized by half of the boson mass, for $\tau$ pairs in simulated $t\bar{t}H$ events (left) and $t\bar{t}Z$ events (right). Only the $\tau^\pm \to \pi^\pm \nu_\tau(\bar{\nu}_\tau)$ decay mode is included in the plots. . . . 172

B.3 Distributions of the energy asymmetry of $\tau^+$ and $\tau^-$ in simulated $t\bar{t}H$ (left) and $t\bar{t}Z$ (right) events in the boson rest frame. One the $\tau^\pm \to \pi^\pm\pi^0\nu_\tau(\bar{\nu}_\tau)$ decay mode in included in the plots. . . . . 172

CHAPTER 1

# INTRODUCTION

In the standard model (SM) of particle physics, the Higgs field is responsible for generating masses for all elementary particles via spontaneous symmetry breaking (SSB). The Higgs boson, which is the quantum excitation of the Higgs field, had been the last missing piece of the standard model until recently. One of the primary motivations to build the Large Hadron Collider (LHC) is to search for the Higgs boson produced in proton-proton collisions. In 2012, the Higgs boson was discovered at the LHC by both ATLAS and CMS experiments [1, 2]. The discovery was based on combinations of all main Higgs production and decay channels in proton-proton collisions at 7 and 8 TeV center-of-mass energies. Since the discovery, extensive studies on properties of the observed Higgs boson have been carried out in order to test its compatibility with the SM prediction, thus ultimately our understanding of electroweak symmetry breaking mechanism.

## 1.1 Higgs boson at the LHC

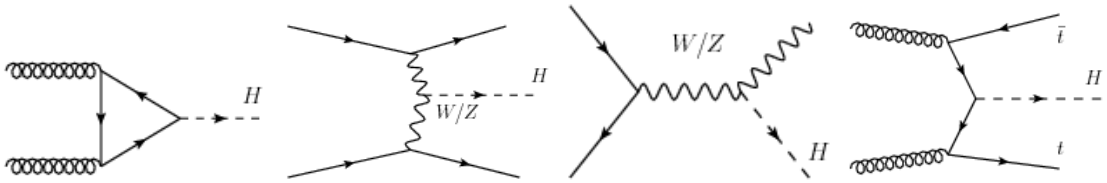


Figure 1.1: Main Higgs boson production modes at the LHC. From left to right: gluon-gluon fusion (ggF), vector boson fusion (VBF), associated Higgs production with vector boson (VH), associated Higgs production with a pair of top quarks ($t\bar{t}H$).

Feynman diagrams of main Higgs boson production modes in proton-proton collisions at the LHC are listed in Fig. 1.1. The production mode with the largest cross section is the gluon-gluon fusion (ggF), followed by the vector boson fusion (VBF). The associated Higgs production with a vector boson (VH) and the associated Higgs production with a pair of top quarks ($t\bar{t}H$) are the other main production modes but with smaller cross sections. The Higgs production cross sections at the LHC as a function of center-of-mass energy, assuming a Higgs mass of 125 GeV, are shown in Fig. 1.2. The center-of-mass energy $E_{CM}$ is expressed in terms of a Mandelstam variable $s = E_{CM}^2$. The Higgs boson pro-

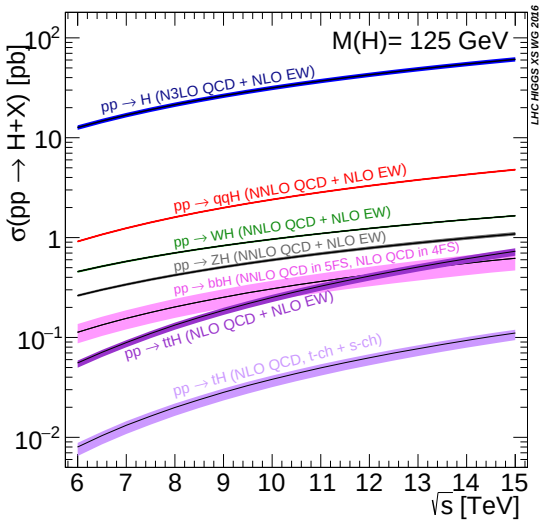


Figure 1.2: The SM Higgs boson production cross sections as a function of the center-of-mass energy for proton-proton collisions [3]. The production mode VBF is indicated as qqH in the plot. In addition to the four main Higgs production modes, associated Higgs productions with a bottom quark pair (bbH) and with a single top quark (tH) are also shown.

duced in proton-proton collisions decays promptly into a pair of gauge bosons or a pair of fermions. Figure 1.3 shows the theoretical Higgs branching ratios of several main decay channels around the Higgs mass of 125 GeV. The bands on the curves in both Fig. 1.2 and Fig. 1.3 indicate theoretical uncertainties.

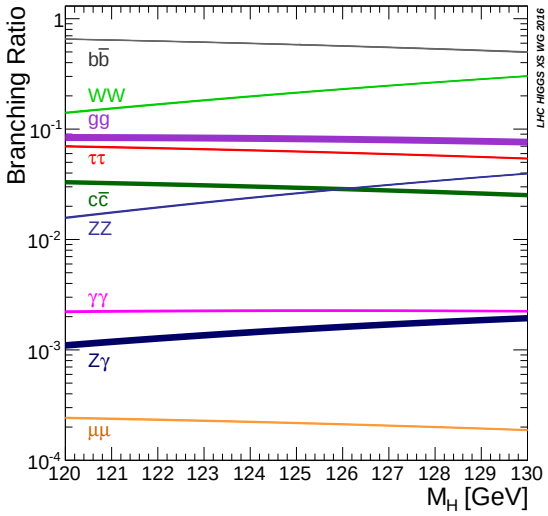


Figure 1.3: Branching ratios of Higgs boson decay modes around $m_H = 125$ GeV [4].

The Higgs boson mass, which had been the last unknown SM parameter at that time, was measured jointly by ATLAS and CMS to be [5]

$$m_H = 125.09 \pm 0.21(\text{stat}) \pm 0.11(\text{syst}) \text{ GeV}, \tag{1.1}$$

Measurements on other properties of the Higgs boson, including production and decay rates [6], spin and parity [7, 8], total width [9–12], and differential cross sections [13–15], are all consistent with the SM predictions within experimental uncertainties.

## 1.2 Associated Higgs boson production with a top quark pair

The Higgs boson production in association with a pair of top quarks, denoted as $t\bar{t}H$, is one of the main Higgs production modes at the LHC and is directly sensitive to the coupling strength between the Higgs boson and the top quark. Figure 1.4 shows the main Feynman diagrams of $t\bar{t}H$ production at the LHC. The

$t\bar{t}H$ process, however, was not observed with $5\sigma$ significance during LHC Run 1 at $\sqrt{s}$ = 7 and 8 TeV due to its small cross section. The center-of-mass proton-proton collision energy was increased to 13 TeV during LHC Run 2 starting from 2015. The $t\bar{t}H$ production mode benefited greatly from the increased collision energy. As shown in Fig. 1.2, the $t\bar{t}H$ cross section increased by about a factor of 4 from $\sqrt{s}$ = 8 TeV to 13 TeV. In 2018, both ATLAS and CMS experiments announced the observation of the $t\bar{t}H$ process using a combination of LHC Run 1 data at $\sqrt{s}$ = 7 and 8 TeV and partial LHC Run 2 data at $\sqrt{s}$ = 13 TeV recorded in 2016 [16, 17].

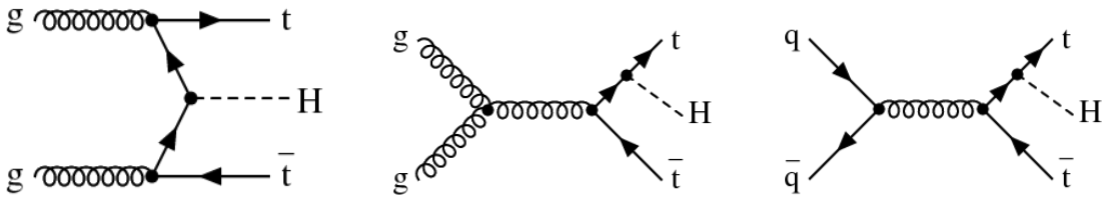


Figure 1.4: Main Feynman diagrams for $t\bar{t}H$ productions at the LHC.

The event signature of the $t\bar{t}H$ process is a superposition of the decay products of the top pair and the Higgs boson. A top quark decays almost always to a W boson and a bottom quark. The W boson can decay either leptonically into an electron (e), muon ($\mu$), or tau ($\tau$) associated with its corresponding neutrino, each with a chance of 11%, or decay hadronically into a pair of light-flavored quarks with a total chance of about 67%. Based on the combination of their decay modes, a pair of top quarks ($t\bar{t}$) can be "leptonic" if both top quarks decay leptonically, or "semi-leptonic" if one decays leptonically and the other hadronically, or "hadronic" if both top quarks decay hadronically. The Higgs boson decays into a pair of gauge bosons or a pair of fermions with different branching ratios, as shown in Fig. 1.3. Combinations of the top pair and Higgs decays give quite complex final states and many analysis channels to consider.

**Leptonic or semi-leptonic** $t\bar{t}$ **+** H $\rightarrow$ bb [18, 19]. In this analysis channel, at least one top quark decays leptonically, and the Higgs boson decays into a pair of bottom quarks. This channel has a relatively high event rate, and the lepton produced from the top decay can be used to trigger such events. The main background process for this channel is the production of a top quark pair associated with heavy-flavored quarks, which is difficult to model accurately in the simulation.

**Hadronic** $t\bar{t}$ **+** H $\rightarrow$ bb [20,21]. This channel targets fully hadronic final states of both top quarks and the Higgs boson and has the highest event rate among all $t\bar{t}$H channels. Dominant backgrounds in this channel are QCD multi-jet events. It also has a sizable $t\bar{t}$+bb background.

**Inclusive** $t\bar{t}$ **+** H $\rightarrow \gamma\gamma$ [22, 23]. The signal rate in this channel is very low due to the small branching ratio of the Higgs boson decaying into a pair of photons. However, both photons from the Higgs decay can be reconstructed if they are within the detector acceptance. The signal extraction is essentially a bump hunt on the diphoton mass distribution. Because of the very clean event signature, results in this analysis channel, despite its low rate, actually drove the sensitivity of the analysis that lead to the observation of the $t\bar{t}$H process.

**Leptonic or semi-leptonic** $t\bar{t}$ **+** H $\rightarrow$ WW, ZZ, $\tau\tau$. This channel, sometimes referred to as $t\bar{t}$H multi-lepton, targets final states with multiple muons, electrons and hadronic taus. Leptons are produced promptly from gauge bosons or tau leptonic decays. Combining the Higgs final states with the leptonic or semi-leptonic $t\bar{t}$ decay, this channel has a rich set of final states with high multiplic-

ities of objects. One example of such event is illustrated in Fig. 1.5. In this

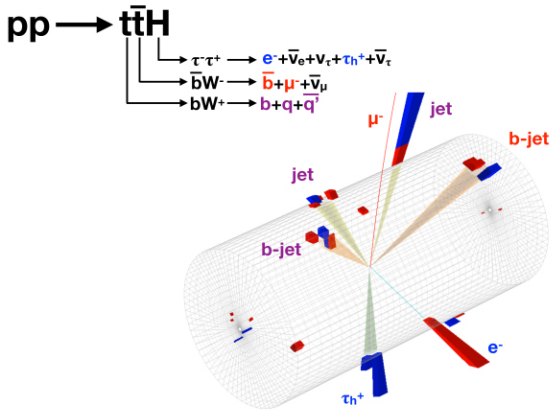


Figure 1.5: A simulated event for tt̄H in multi-lepton final states [24].

particular event, the top quark decays into a bottom quark and a $W^+$, which decays hadronically into a pair of light-flavored quarks. The anti-top decays into an anti-bottom quark and $W^-$, which decays leptonically into a muon and a muon anti-neutrino. The Higgs boson in this event decays into a pair of tau leptons, one of which decays into an electron, an electron anti-neutrino, and a tau neutrino. The other tau lepton decays hadronically, denoted as $\tau_h$. All neutrinos produced in the decay escape the detection of the experiment and would be partially reconstructed as missing transverse momentum. This channel has relatively lower signal rates compared to the hadronic channels, but also has a relatively cleaner event signature due to leptons in the final state. Results of ATLAS and CMS analyses in this channel based on early LHC Run 2 data at $\sqrt{s} = 13$ TeV are reported in Refs. [25] and [26]. More detailed discussions and up-to-date information on the tt̄H multi-lepton analysis with the CMS experiment at 13 TeV are presented in the rest of this thesis.

CHAPTER 2

# STANDARD MODEL AND HIGGS COUPLINGS

This chapter briefly overviews the ingredients in the standard model Lagrangian density, including boson and fermion fields in Section 2.1, the Higgs field and the spontaneous symmetry breaking in Section 2.2, and Yukawa couplings in Section 2.3. The notation and derivation in these sections generally follow Ref. [27]. Measurements on Yukawa coupling strengths at the LHC, particularly the top-Higgs coupling, are discussed in Section 2.4.

## 2.1 Boson and fermion fields

The standard model of particle physics encapsulates our understanding of the fundamental matter particles and their interaction. Matter particles include quarks and leptons, as well as their anti-particles. Both quarks and leptons are fermions of spin 1/2. Interactions between matters particles are mediated by gauge bosons. Three of the four fundamental forces are described in the SM: the strong force, carried by gluons; the weak force, carried by W and Z bosons; the electromagnetic force, carried by photons. The gravitational force is not included, but its impact on particles is negligible in laboratory experiments with energies that are currently accessible. The SM is remarkably successfully so far in explaining nearly all experimental data in particle physics and predicting preciously many phenomena that have been verified by experiments.

The SM describes particles and their interaction via a gauge quantum field

theory based on the gauge group

$$SU(3) \times SU(2) \times U(1). \tag{2.1}$$

The $SU(3)$ factor is the gauge group for quantum chromodynamics (QCD) describing the strong force. In QCD, there are eight massless gauge bosons $G^i, i = 1 \cdots 8$, i.e. gluons that couple to matter particles with color charges i.e. quarks. The weak and electromagnetic forces are combined in a unified electroweak gauge theory based on the gauge group $SU(2) \times U(1)$. The $SU(2)$ group is associated with three gauge bosons $W^i, i = 1, 2, 3$, and only acts to the left chiral part of fermions. The left handedness of the electroweak gauge group is often indicated by adding a subscript $L$ to $SU(2)$. The $U(1)$ group acts on both left- and right-chiral fermions but with different gauge couplings. It has one gauge boson, denoted as $B$. The $SU(2)$ gauge bosons $W^1$ and $W^2$ can be expressed as linear combinations of the "raising" and "lower" operators $W^\pm$ of $SU(2)$. In addition, a linear combination of $W^3$ and $B$ incorporates the photon field of quantum electrodynamics, and the orthogonal combination yields $Z$ boson field. After the spontaneous symmetry breaking, which will be discussed in Section 2.2, the photon remains massless while $W^\pm$ and $Z$ acquire masses. The Lagrangian density of the gauge sector can be collectively written as

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} F^i_{\mu\nu} F^{\mu\nu i}, \tag{2.2}$$

where $F^i_{\mu\nu}$ represents any field strength tensor for $SU(3)$, $SU(2)$, and $U(1)$ gauge bosons.

$$F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\nu A^i_\mu - g c_{ijk} A^j_\mu A^k_\nu \tag{2.3}$$

Equation 2.3 is a generic expression of the field strength, where $A^i$ is a boson field ($G^i, i = 1 \cdots 8$ in case of $SU(3)$, $W^i, i = 1, 2, 3$ in case of $SU(2)$, and $B$ in case

of $U(1)$), $g$ parameterize the gauge coupling strength, and $c_{ijk}$ are the structure constants of the corresponding group. In case of $U(1)$, $c_{ijk} = 0$.

The fermion sector of the SM includes three generations of quarks and leptons. Each generation of quarks or leptons has an $SU(2)$ invariant left-chiral doublet field. Each component of the doublet has a right-chiral counterpart that is a $SU(2)$ singlet. The field contents for quarks and leptons can be expressed as

$$q_{mL} = \begin{pmatrix} u_{mL} \\ d_{mL} \end{pmatrix}, \; u_{mR}, \; d_{mR}, \tag{2.4}$$

$$\ell_{mL} = \begin{pmatrix} \nu_{mL} \\ e_{mL} \end{pmatrix}, \; e_{mR}, \; (\nu_{mR}), \tag{2.5}$$

where $m = 1, 2, 3$ labels three generations, and $L$ and $R$ refer to left and right handedness, respectively. The up-type $u$ (down-type $d$) quark has an electric charge of $2/3$ ($-1/3$) and three types of color charges. The lepton $e$ has an electric charge of -1 and does not carry any color charge. The neutrino $\nu$ is both color and electrically neutral. The right-handed neutrino $\nu_R$, which is motivated by many models for neutrino mass, is put in the parenthesis in Eq. 2.5 due to lack of experimental evidences so far of its existence. The representation of the SM gauge group $SU(3) \times SU(2) \times U(1)$ for fermion fields can be denoted as $(n_3, n_2, y)$, in which $n_3$ and $n_2$ are representations of $SU(3)$ and $SU(2)$, respectively. The symbol $y$ is the weak hypercharge defined as $y = q - t_3$, where $q$ is the electric charge and $t_3$ is the eigenvalue of the third generator of $SU(2)$. Representations of the SM fermion fields are summarized in Table 2.1. The Lagrangian density of the fermion sector can be expressed as

$$\mathcal{L}_{\mathrm{f}} = i\bar{\psi}\gamma^{\mu}\mathcal{D}_{\mu}\psi \tag{2.6}$$

where $\psi$ represents any fermion field, $\mathcal{D}_{\mu}$ is its corresponding covariant deriva-

Table 2.1: Representations of SM gauge group for fermions.

| Fermion field | Representation |
|---|---|
| $q_L$ | $(3, 2, 1/6)$ |
| $u_R$ | $(3, 1, 2/3)$ |
| $d_R$ | $(3, 1, -1/3)$ |
| $\ell_L$ | $(1, 2, -1/2)$ |
| $e_R$ | $(1, 1, -1)$ |

tive. A generic expression of the covariant derivative is given in Eq. 2.7.

$$\mathcal{D}_\mu = \partial_\mu + i \sum g \vec{T} \cdot \vec{A}_\mu \tag{2.7}$$

The summation includes all relevant groups of the fermion according to its representation given in Table 2.1. For each gauge group, $g$ is the gauge coupling, $\vec{A}_\mu$ is a vector of the gauge boson fields, and $\vec{T}$ are generators in the defining representation of the group. For $SU(2)$ and $SU(3)$, representation matrices of their generators are Pauli and Gell-Mann matrices divided by two, respectively.

## 2.2 The Higgs mechanism

The SM gauge symmetry does not allow bare mass terms in either $\mathcal{L}_{\text{gauge}}$ for gauge bosons or $\mathcal{L}_{\text{f}}$ for fermions. Effective masses of these particles, however, can be generated via the spontaneous symmetry breaking by introducing a Higgs field. If the Higgs field has a nonzero vacuum expectation value (VEV), the electroweak gauge symmetry is broken in the ground state.

A complex scalar Higgs doublet that transforms under $SU(2) \times U(1)$ electroweak gauge symmetry can be written as

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.8}$$

The potential of the field can only be expressed in the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2. \tag{2.9}$$

in order to be renormalizable and $SU(2) \times U(1)$ invariant. In Eq. 2.9, $\lambda$ must be greater than zero so that $V$ is bounded from below. If $\mu^2 < 0$, the minimum of the potential is no longer at $\phi = (0,0)^{\mathrm{T}}$. Instead, the minimum is at $\phi^\dagger \phi = v^2/2$, where

$$v = \sqrt{\frac{-\mu^2}{\lambda}} > 0. \tag{2.10}$$

A ground state of the Higgs field can therefore be written as[1]

$$\phi_0 \equiv \langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.11}$$

The nonzero VEV $v$ breaks spontaneously the $SU(2) \times U(1)$ electroweak symmetry in the ground state of the Higgs field. To generate masses for the weak gauge bosons, the Higgs field is quantized around the classical vacuum, i.e. the field is expanded around its ground state:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \tag{2.12}$$

In Eq. 2.12, $H$ is a Hermitian scalar field and will be the physical Higgs boson.

The SM Lagrangian density of the Higgs sector is

$$\mathcal{L}_\phi = (\mathcal{D}_\mu \phi)^\dagger (\mathcal{D}^\mu \phi) - V(\phi) \tag{2.13}$$

In the above equation, the Higgs covariant derivative is

$$\mathcal{D}_\mu = \partial_\mu - i\frac{g}{2}\sigma^i W_\mu^i - i\frac{g'}{2} B_\mu, \tag{2.14}$$

[1]Other ground states can be obtained by an $SU(2) \times U(1)$ rotation.

where $\sigma^i$ are Pauli matrices, and $g$ and $g'$ are $SU(2)$ and $U(1)$ gauge couplings, respectively. The second term in the covariant derivative can be explicitly written as

$$i\frac{g}{2}\sigma^i W_\mu^i = i\frac{g}{2}\begin{pmatrix} W_\mu^3 & W_\mu^1 + iW_\mu^2 \\ W_\mu^1 - iW_\mu^2 & -W_\mu^3 \end{pmatrix} = i\frac{g}{2}\begin{pmatrix} W_\mu^3 & \sqrt{2}W_\mu^- \\ \sqrt{2}W_\mu^+ & -W_\mu^3 \end{pmatrix}, \tag{2.15}$$

where $W_\mu^\pm \equiv (W_\mu^1 \mp iW_\mu^2)/\sqrt{2}$.

The Higgs kinematic term is thus

$$\begin{aligned} (\mathcal{D}_\mu\phi)^\dagger(\mathcal{D}^\mu\phi) &= \frac{1}{2}(0\ v)\left[\frac{g}{2}\sigma^i W_\mu^i + \frac{g'}{2}B_\mu\right]^2\begin{pmatrix} 0 \\ v \end{pmatrix} + H\text{ terms} && (2.16)\\ &= \frac{1}{8}(0\ v)\begin{pmatrix} g'B_\mu + gW_\mu^3 & \sqrt{2}gW_\mu^- \\ \sqrt{2}gW_\mu^+ & g'B_\mu - gW_\mu^3 \end{pmatrix}^2\begin{pmatrix} 0 \\ v \end{pmatrix} + H\text{ terms} && (2.17)\\ &= \frac{v^2g^2}{4}W_\mu^+W^{-\mu} + \frac{v^2}{8}(g'B_\mu - gW_\mu^3)^2 + H\text{ terms} && (2.18)\end{aligned}$$

By defining

$$Z_\mu \equiv \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2+g'^2}}, \tag{2.19}$$

the Higgs kinematic term can be simplified as

$$(\mathcal{D}_\mu\phi)^\dagger(\mathcal{D}^\mu\phi) = M_W^2W_\mu^+W^{-\mu} + \frac{1}{2}M_Z^2Z_\mu Z^\mu + H\text{ terms}, \tag{2.20}$$

where

$$M_W \equiv \frac{vg}{2}\text{ and }M_Z \equiv \frac{v}{2}\sqrt{g^2+g'^2} \tag{2.21}$$

are W and Z boson masses, respectively.

The $H$ terms in the rest of the Lagrangian, which include interaction terms of Higgs and gauge bosons, can also be expressed in terms of $W_\mu^\pm$, $Z_\mu$, and their

masses $M_W$, $M_Z$. The full Lagrangian density of the Higgs sector after the SSB is

$$\begin{aligned}\mathcal{L}_\phi = & M_W^2 W_\mu^+ W^{-\mu}\left(1+\frac{H}{v}\right)^2 + \frac{1}{2}M_Z^2 Z_\mu Z^\mu\left(1+\frac{H}{v}\right)^2 \\ & + \frac{1}{2}(\partial_\mu H)^2 + \mu^2 H^2 + \frac{\mu^2}{v}H^3 + \frac{\mu^2}{4v^2}H^4 - \frac{1}{4}\mu^2 v^2.\end{aligned} \tag{2.22}$$

The tree-level mass for the the Higgs boson can be read off from Eq. 2.22:

$$M_H = \sqrt{-2\mu^2} = \sqrt{2\lambda}v. \tag{2.23}$$

## 2.3 Yukawa couplings

Fermions also acquire masses by the SSB. In order for the SM to incorporate fermion masses, Yukawa coupling terms between the Higgs boson and fermions are added into the SM Lagrangian. The Yukawa coupling terms are required to be both Lorentz and $SU(2)\times U(1)$ gauge invariant before the SSB.

For a lepton, for example, such Yukawa coupling term can be written as[2]

$$\mathcal{L}_{\text{Yuk},e} = -y_e \bar{\ell}_L \phi e_R + h.c. \tag{2.24}$$

where $y_e$ is its Yukawa coupling strength, and $h.c.$ stands for Hermitian conjugate. After the SSB, the Higgs field $\phi$ is chosen as in Eq. 2.12. The Yukawa term for the lepton becomes

$$\begin{aligned}\mathcal{L}_{\text{Yuk},e} &= -\frac{y_e}{\sqrt{2}}\,(\bar{\nu}_L\,\bar{e}_L)\begin{pmatrix}0\\ \\ v+H\end{pmatrix}e_R + h.c. \\ &= -\frac{y_e v}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{y_e}{\sqrt{2}}H(\bar{e}_L e_R + \bar{e}_R e_L) \\ &= -m_e \bar{e}e - \frac{m_e}{v}H\bar{e}e,\end{aligned} \tag{2.25}$$

[2] Assume no right-handed neutrino $\nu_R$.

where $m_e = (y_e v)/\sqrt{2}$ is the lepton mass.

Quarks acquire masses in a similar way after the SSB by coupling to the Higgs field. The Yukawa coupling term for quarks, assuming three generations, can be written as

$$\mathcal{L}_{\text{Yuk,quarks}} = -\sum_{m,n=1}^{3} \left( y^u_{mn} \bar{q}_{mL} \tilde{\phi} u_{nR} + y^d_{mn} \bar{q}_{mL} \phi d_{nR} \right) + h.c. \tag{2.26}$$

where $\tilde{\phi}$ is a conjugate form of $\phi$ introduced to deal with the right-handed component of the up-type quark:

$$\tilde{\phi} \equiv i\sigma^2 \phi^\dagger = \begin{pmatrix} \phi^{0\dagger} \\ -\phi^- \end{pmatrix} \tag{2.27}$$

The Yukawa coupling matrix $y^u_{mn}$ and $y^d_{mn}$ encode the quark masses and mixing among flavors, and do not need to be diagonal. It is also worth mentioning that quark fields in Eq. 2.26 are weak eigenstates. A set of unitary transformation matrices can be introduced to diagonalize the Yukawa coupling matrices and transform both the left- and right-handed components into mass eigenstates. After the SSB, the Yukawa coupling terms for quarks can be expressed as

$$\mathcal{L}_{\text{Yuk,quarks}} = -\left(1 + \frac{H}{v}\right) \sum_{j=1}^{3} \left( m^u_j \bar{u}'_j u'_j + m^d_j \bar{d}'_j d'_j \right) \tag{2.28}$$

The prime symbols in superscripts indicate the quark fields are mass eigenstates. The index $j$ ranges from 1 to 3, corresponding to three generations of quarks in the SM. As shown in both Eq. 2.25 and Eq. 2.28, the Yukawa coupling between a fermion and the Higgs boson is proportional to the fermion mass.

## 2.4 Measure the top-Higgs coupling

Measuring strengths of Higgs coupling to other particles and comparing them to those predicted by the SM are crucial for testing our understanding of electroweak symmetry breaking. It also helps to probe new physics beyond the standard model (BSM). Despite its remarkable success so far, the SM is generally believed to be a low-energy approximation of a more fundamental theory. Besides its unnaturalness (e.g. Hierarchy problem, Strong CP problem), the SM does not explain e.g. the nature of dark matter and the matter-antimatter asymmetry. Many BSM theories, which attempt to address the above problems, also modify the SM electroweak symmetry breaking mechanism, introducing deviations of Higgs couplings with respect to the SM predictions.

Among all Higgs couplings, the top-Higgs coupling is of particular interest. As the heaviest SM elementary particle, the top quark has the strongest coupling to the Higgs boson and is mainly responsible for the instability of Higgs mass against radiative corrections. New physics at a higher energy scale concerning electroweak symmetry breaking can affect the low energy domain as an effective field theory via dimension-6 operators, which may preferentially couple to the top quark. As a result, the effective top-Higgs coupling can deviate from the SM prediction. By measuring the top-Higgs coupling precisely, quantitative conclusions can be drawn on these higher dimension operators. In addition, the top Yukawa coupling involved in loop processes are also sensitive to the existence of BSM physics. The top-Higgs coupling strength can be extracted either indirectly from the gluon-gluon fusion process of the Higgs production (c.f. leftmost Feynman diagram in Fig. 1.1), assuming no BSM contribution in the loop, or directly from the tree-level $t\bar{t}H$ process (c.f. rightmost diagram in

Fig. 1.1). Any inconsistency between the indirect and direct top-Higgs coupling measurements would hint the existence of new physics beyond the standard model.

### 2.4.1 Parameterization

Yukawa couplings are extracted from Higgs production cross section and branch ratio measurements. A few assumptions are made in order to extract the Higgs coupling strengths:

- The Higgs boson is a CP-even scalar. BSM couplings only modify the coupling strength but not the tensor structure.
- Other resonances that may overlap with the observed Higgs boson are not considered.
- The total width of the Higgs boson near its mass of 125 GeV is narrow, so its production and decay processes can be factorized.

For a given process $i \to \mathrm{H} \to f$, the event yield $N$ is proportional to the cross section and branching ratio:

$$N \propto \sigma_i \cdot \mathrm{B}^f = \frac{\sigma_i \Gamma_f}{\Gamma_{\mathrm{tot}}} \tag{2.29}$$

In the above expression, $\sigma_i$ is the cross section of the production mode $i \to \mathrm{H}$; $\mathrm{B}^f$ is the branching ratio of the decay channel $\mathrm{H} \to f$. The symbol $\Gamma_f$ and $\Gamma_{\mathrm{tot}}$ are the partial decay width for $\mathrm{H} \to f$ and the total width of the Higgs boson, respectively. By grouping the relevant Higgs production and decay processes, the Higgs coupling strengths are parameterized as ratios of the observed cross

section or decay width to that predicted by the SM in a so-called $\kappa$-framework. The coupling modifier $\kappa$ for a Yukawa coupling $j$ is defined as

$$\kappa_j^2 = \frac{\sigma_j}{\sigma_j^{\mathrm{SM}}} \quad \text{or} \quad \kappa_j^2 = \frac{\Gamma_j}{\Gamma_j^{\mathrm{SM}}} \tag{2.30}$$

### 2.4.2 Indirect measurements

For a loop-induced process involving top quarks, such as gluon-gluon fusion of the Higgs production ($gg \to \mathrm{H}$) or Higgs decaying into a pair of photons ($\mathrm{H} \to \gamma\gamma$), either an effective coupling modifier, $\kappa_g$ for gluons and $\kappa_\gamma$ for photons, or a combination of tree-level coupling modifiers of the SM particles that contribute to the loop is used to parameterize the process. In the latter case, the effective coupling modifier for a ggF process, assuming no BSM contributions, can be expressed as

$$\kappa_g^2 \sim 1.06\kappa_{\mathrm{t}}^2 + 0.01\kappa_{\mathrm{b}}^2 - 0.07\kappa_{\mathrm{t}}\kappa_{\mathrm{b}}. \tag{2.31}$$

Similarly for $\mathrm{H} \to \gamma\gamma$, the effective coupling modifier can be written as

$$\kappa_\gamma^2 \sim 1.59\kappa_{\mathrm{W}}^2 + 0.07\kappa_{\mathrm{t}}^2 - 0.66\kappa_{\mathrm{W}}\kappa_{\mathrm{t}}. \tag{2.32}$$

In Eq. 2.31 and Eq. 2.32, $\kappa_{\mathrm{t}}$, $\kappa_{\mathrm{b}}$, and $\kappa_{\mathrm{W}}$ are Yukawa coupling modifiers for top quarks, bottom quarks, and W bosons, respectively.

The $\kappa_{\mathrm{t}}$ can be extracted from these loop processes, mostly from the ggF process in which top quarks have the dominant contribution, as shown in Eq. 2.31. In order to properly handle correlations among all Yukawa couplings involved in both the production and decay processes and also to take advantage of Higgs measurements in all analysis channels, all relevant $\kappa$'s under a chosen parameterization are extracted simultaneously from a fit. One such parameterization

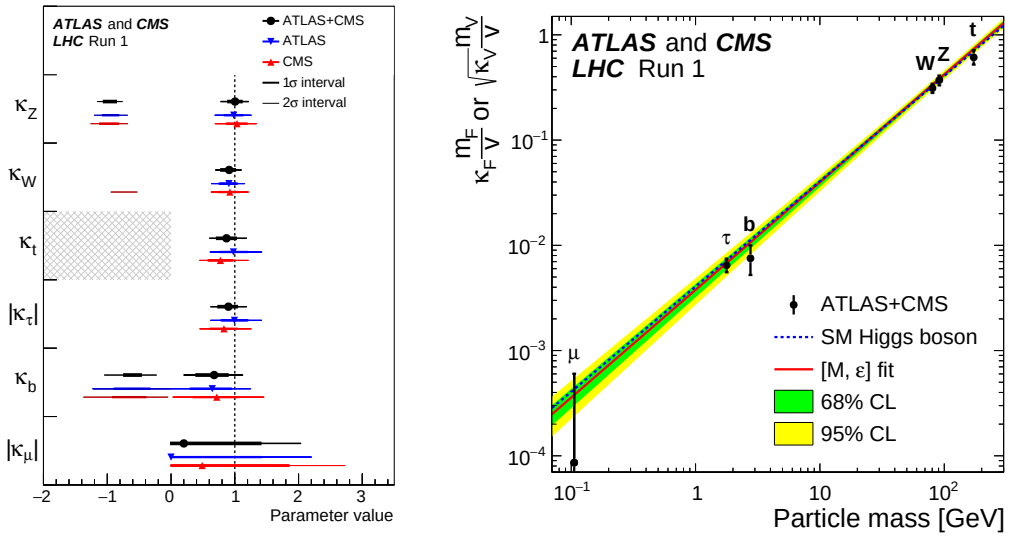


Figure 2.1: Left: best fit values of Higgs coupling modifiers in the 6-$\kappa$-parameter framework for ATLAS (blue) and CMS (red) individually and for the combination of data from the two experiments (black). It is assumed that no BSM particles contributed to the loop processes. Right: best fit values of Yukawa coupling strengths as a function of particle masses based the combined results of ATLAS and CMS. The SM prediction is indicated with the blue dashed line. Data are also fitted with a linear function as indicated by the red solid line with 68% and 95% confidence level (CL) bands. Plots are taken from Ref. [6].

consists of six free tree-level coupling modifiers: $\kappa_W$, $\kappa_Z$, $\kappa_t$, $\kappa_b$, $\kappa_\tau$, and $\kappa_\mu$. The effective coupling modifiers for $gg \to H$ and $H \to \gamma\gamma$ loop-induced processes are expressed in terms of the tree-level coupling modifiers and their interference, as in Eq. 2.31 and 2.32. Results from ATLAS and CMS alone, as well as the combined result, using LHC data at $\sqrt{s} = 7$ and 8 TeV ("LHC Run 1") are summarized in the left plot of Fig. 2.1. The ATLAS and CMS combined result for the top-Higgs coupling modifier $\kappa_t$ is $0.87^{+0.15}_{-0.15}$ [6].

The combined results are also converted to the Yukawa coupling strength $y$ to better illustrate the agreement between the measured value and the SM predictions. The parameter $y$ is defined as

$$y_V = \sqrt{\kappa_V}\frac{m_V}{v} \quad \text{and} \quad y_F = \kappa_F \frac{m_F}{v}, \tag{2.33}$$

where subscripts V and F denote vector bosons and fermions, respectively, and $v$ is the VEV of the Higgs field. The measured Yukawa coupling strengths as a function of particle masses are compared to the SM predictions in the right plot of Fig. 2.1.

### 2.4.3 Direct measurements

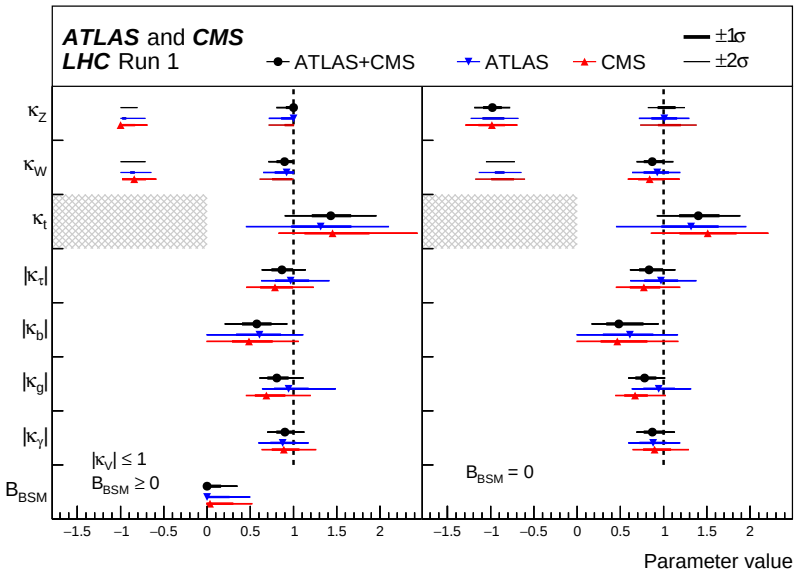


Figure 2.2: Fit results of Yukawa coupling modifiers with $\kappa_t$ directly constrained by the $t\bar{t}H$ process under two parameterizations: one allows non-zero BSM Higgs branching ratio ($|\kappa_V| \leq 1$ and $B_{BSM} \geq 0$); the other does not ($B_{BSM} = 0$) [6]. Results based on ATLAS (blue) and CMS (red) data separately, as well as on the combined data (black), are shown in the plots.

A different choice of the Higgs coupling modifier parameterization can be used to extract the top-Higgs coupling from the tree-level $t\bar{t}H$ process. In this parameterization, $\kappa_t$ is directly constrained from measurements of the $t\bar{t}H$ process, and two effective coupling modifiers $\kappa_g$ and $\kappa_\gamma$ are added as free parameters. Other free $\kappa$ parameters include $\kappa_W$, $\kappa_Z$, $\kappa_b$, and $\kappa_\tau$ for W bosons, Z bosons,

bottom quarks, and $\tau$ leptons, respectively. Generation universality is assumed for Yukawa couplings of quarks and leptons. Two scenarios are considered for the fit. One allows the possibility of the Higgs boson decaying into BSM particles and adds a non-negative Higgs BSM branching ratio $B_{BSM}$ as the eighth free parameter to the fit. It also assumes that $\kappa_W$ and $\kappa_Z$ are of the same sign and that both $\kappa_W$ and $\kappa_Z$ satisfy $|\kappa_V| < 1$. The other scenario sets $B_{BSM}$ to zero.

The combined results of ATLAS and CMS, as well as the individual results, are shown in Fig. 2.2 for both scenarios using LHC Run 1 data. The measured top-Higgs coupling modifier is $\kappa_t = 1.43^{+0.23}_{-0.22}$ [6] in the parameterization assuming $B_{BSM} \geq 0$ and $|\kappa_V| \leq 1$. In case of the scenario with $B_{BSM} = 0$, the measured $\kappa_t$ is $1.40^{+0.24}_{-0.21}$ [6]. Uncertainties on $\kappa_t$ in both scenarios are relatively large, because the analysis sensitivity in $t\bar{t}H$ process is statistically limited during LHC Run 1.

CHAPTER 3

# EXPERIMENTAL APPARATUS

This chapter describes the experimental apparatus used in the $t\bar{t}H$ analysis. The analysis is conducted by studying the proton-proton collision events that are produced at the CERN Large Hadron Collider and recorded by the Compact Muon Solenoid (CMS) detector. A brief overview of the LHC is given in Section 3.1. The CMS detector and its subsystems are described in Section 3.2.

## 3.1 Large Hadron Collider

The LHC [28] is a circular hadron collider with superconducting magnets at CERN near Geneva, Switzerland. It is located in the previous LEP tunnel with a circumference of approximately 27 km and about 100 m underground. The LHC can be configured to collide either protons or heavy ions. The designed center-of-mass energy for proton collisions is 14 TeV and the design peak luminosity is $10^{34}$ $cm^{-2}$ $s^{-1}$. During Run 1 (2010 - 2012), the LHC operated at collision energies of 7 and 8 TeV. The data used in the analysis presented in this thesis were taken during LHC Run 2 (2015 - 2018) with a center-of-mass energy of 13 TeV.

In the LHC, two particle beams with the same charge travel in opposite directions in separate ultrahigh vacuum ($10^{-10}$ mbar) tubes before they are made to collide. Given the radius of the ring, the maximum energy that can be achieved is proportional to the bending magnetic filed. The LHC uses a total of 9593 magnets, among which 1232 are the main dipole magnets creating an

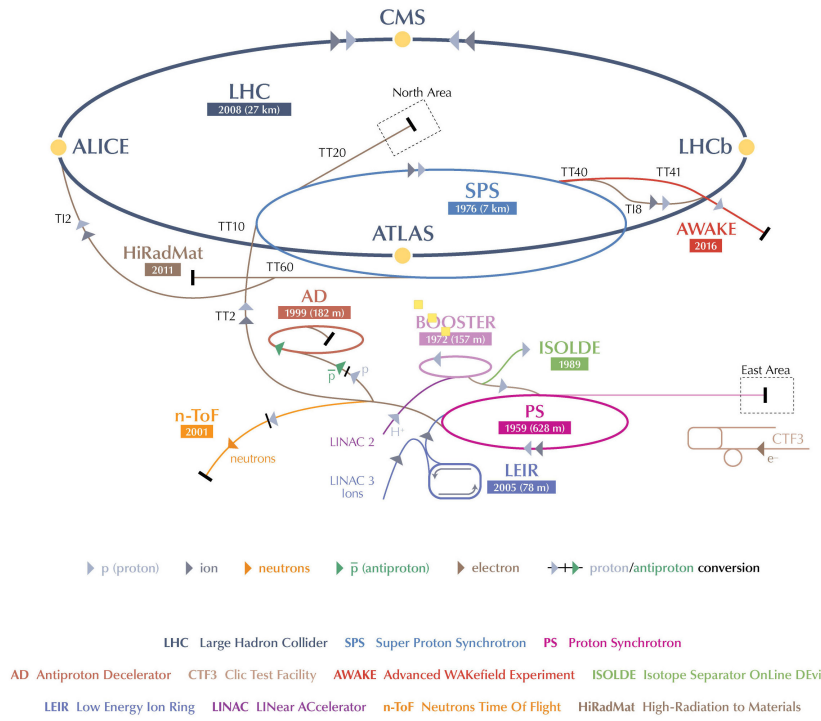


Figure 3.1: The accelerator complex at CERN including the LHC [29]. The diagram also shows the four major experiments on the LHC.

8.33 T magnetic field to maintain the circular trajectory of the beams. There are also 392 main quadrupole magnets that are used to focus the beams. The LHC magnets use superconducting niobium-titanium (NbTi) cables that are cooled by superfluid helium to a temperature of 1.9 K.

As the most powerful accelerator in the world, the LHC is the last and the largest component in the chain of the CERN accelerator complex. The accelerator complex is illustrated in Fig. 3.1. The protons are produced by stripping electrons from hydrogen atoms with an electric field. They are accelerated by a linear accelerator, Linac2, to an energy of 50 MeV, before they are injected into the Proton Synchrotron Booster (PSB). The PSB accelerates the protons to 1.4 GeV and feeds them to the Proton Synchrotron (PS),[1] where the beam energy is

[1]In case of heavy ion running mode, the PS is supplied by the Low Energy Ion Ring (LEIR) instead.

further increased to 25 GeV. The protons are then transferred to the Super Proton Synchrotron (SPS). This is the last stage before the LHC and the protons are accelerated to 450 GeV. Finally, the proton beams are sent to the LHC ring where each beam is accelerated to the maximum energy, currently 6.5 TeV. Each proton beam contains maximum 2556 bunches with $1.2 \times 10^{11}$ protons per bunch. The bunch spacing is 25 ns, or equivalently the proton bunches collide at each interaction point at a rate of 40 MHz.[2] On average each bunch crossing has about 20 overlapping proton-proton interactions (pileup) during LHC Run 1 and about 30 during LHC Run 2.

There are four main experiments located on the LHC ring: ALICE [30], ATLAS [31], CMS [32], and LHCb [33]. ATLAS and CMS are two multi-purpose detectors positioned on diametrically opposite sides of the LHC ring. They have similar physics goals and cover a wide range of physics programs at the LHC, but are based on different designs and technologies. ALICE specializes in measuring heavy ion collisions and studying quark-gluon plasma. LHCb is designed to study B physics and CP violation.

## 3.2 Compact Muon Solenoid detector

The CMS detector is one of the two multi-purpose particle detectors at the LHC. It weighs 14000 tons and is a cylindrical detector with a length of 28.7 m and a diameter of 15.0 m. A distinctive feature of the CMS detector is its superconducting solenoid, providing a 3.8 T magnetic field along the beam direction. The CMS apparatus consists of several concentric layers of subdetectors,

[2] During most of the LHC Run 1, the bunch spacing was doubled to 50 ns.

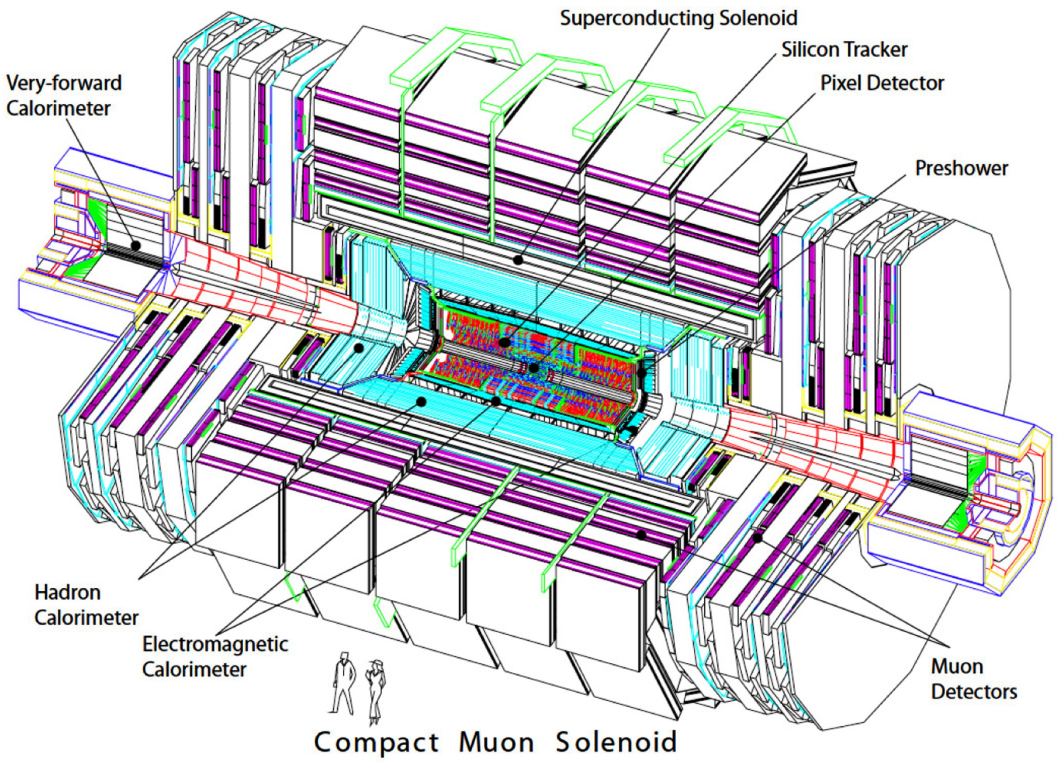


Figure 3.2: The overall layout of CMS [32]. Detector components from innermost to outermost: silicon tracker, electromagnetic calorimeter, hadron calorimeter, superconducting solenoid, and muon detector.

shown in Fig. 3.2, including a silicon pixel and strip tracker, an electromagnetic calorimeter (ECAL), a hadron calorimeter (HCAL), and a muon system. The CMS subsystems are described further in the following sections. More detailed information can be found in Ref. [32]. Each subsystem plays a particular role in reconstructing either trajectory or energy of the particles produced from the proton collisions. The reconstruction of physics objects are discussed in Chapter 4.

### 3.2.1 Coordinate system

The coordinate system adopted by CMS has the origin centered at the nominal collision point inside the detector. The $x$-axis points towards the center of the LHC ring. The $y$-axis points upward perpendicular to the plane of the LHC. The $z$-axis is along the beam pipe with its direction determined by the right-hand rule. It is often more convenient to use cylindrical coordinates due to the shape of the detector. The radial coordinate $r$ is defined in the $x$-$y$ plane, and the azimuthal angle $\phi$ is measured from the $x$-axis in the transverse plane. The polar angle $\theta$ is measured in the $r$-$z$ plane from the positive $z$-axis. A more useful alternative in collider physics, pseudorapidity, is defined based on the polar angle as

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right), \tag{3.1}$$

where $\vec{p}$ and $p_z$ is the three-momentum and its $z$ component, respectively. In case of highly relativistic particles, the pseudorapidity becomes rapidity

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.2}$$

where $E$ is the particle's energy. Differences in rapidity are invariant under longitudinal Lorentz boosts. Angular separations between two particles in terms of $\Delta\eta$ and $\Delta\phi$ therefore do not depend on the choice of reference frame along the beam axis.

### 3.2.2 Solenoid

The CMS superconducting solenoid provides a uniform magnetic field of 3.8 T along the $z$-axis. It is 12.5 m long and has a diameter of 6 m. All CMS subdetectors except the muon system and a thin layer of hadron calorimeter are within

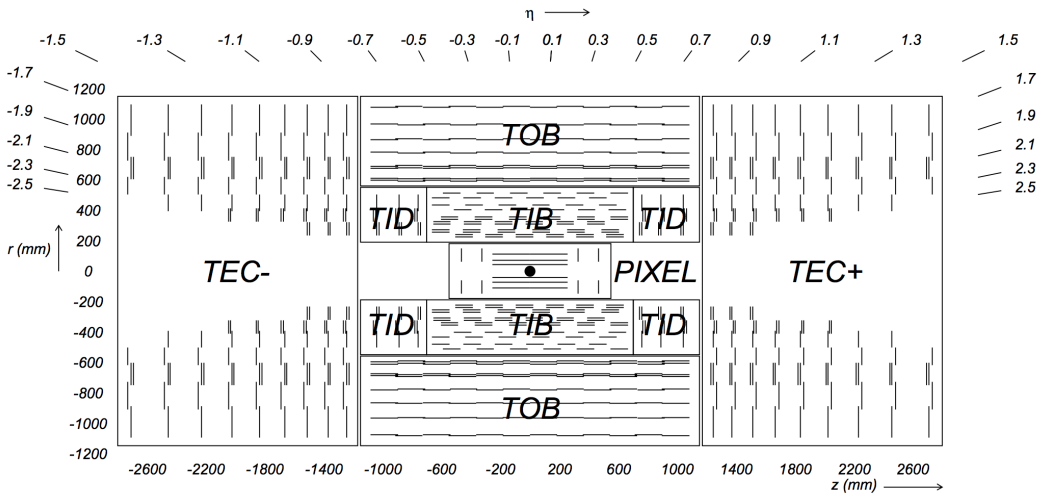


Figure 3.3: Schematic cross section through the CMS tracker in the $r$-$z$ plane [32]. Each sensor module is represented by a line segment.

the volume of the solenoid. The solenoid coil is comprised of 4 layers of NbTi conductor reinforced with aluminum alloy. Liquid helium is used to cool down the solenoid to about 4 K. The flux is returned through an iron yoke composed of 5 barrel wheels and 3 endcap disks on each side of the solenoid.

### 3.2.3 Silicon tracker

The CMS inner tracking system [34] is based on silicon detector technology featuring high granularity and fast response. The CMS tracker is designed to reconstruct charged particle trajectories in the homogeneous magnetic field provided by the superconducting solenoid. The primary vertices of the proton collisions as well as the secondary decay vertices are also reconstructed by the tracker. The CMS tracker consists of a pixel detector and a silicon strip detector, covering the full $\phi$ range and $|\eta| < 2.5$. A schematic of the CMS tracker in the $r$-$z$ plane is shown in Fig. 3.3.

**Pixel Detector**

The pixel detector is the detector closest to the beam pipe. It is designed to provide precise 3D measurements of tracking points in order to have good vertex reconstruction. The pixel detector has three barrel layers (BPix) at mean radii of 4.4, 7.3, and 10.2 cm with two endcap disks on both sides (FPix) at $z = \pm 34.5$ and $z = \pm 46.6$ cm. This arrangement ensures each track within the $\eta$ range has at least 3 pixel hits.

BPix and FPix contain in total 48 million and 18 million pixels respectively. Each pixel cell has a size of $100 \times 150$ $\mu$m$^2$, providing comparable track resolution in both $r$-$\phi$ and $z$ directions. The size of the pixel is driven by the physics requirement on impact parameter resolution. The pixel sensors adopt an n+ pixel on n-substrate design. Sensor signals are read out by the Read-out chips (ROCs), each of which is bump bonded to $52 \times 80$ pixels. A group of ROCs (8 or 16 in BPix, 21 or 24 in FPix) are controlled by a Token Bit Manager (TBM) chip that sends data via an optical link to the Front End Driver (FED).

As the closest subdetector to the interaction region, the pixel detector suffers the highest radiation damage. The original pixel detector was replaced during the extended year-end technical stop between 2016 and 2017. The upgraded Phase-1 pixel detector [35] is designed to maintain a good tracking performance at luminosities up to $2 \times 10^{34}$ cm$^{-2}$ s$^{-1}$ and average pileup up to 50. It has three disks in each endcap and four (instead of three) barrel layers, providing an additional fourth pixel hits within the tracking $\eta$ range. A new cooling system using $CO_2$ was installed, and the overall material budget in the tracker volume is reduced.

**Strip Tracker**

The reduced particle flux at larger radii allows the strip tracker to use silicon micro-strip sensors. The sensors are single sided and are of p-on-n type. The strip pitch is around 100 $\mu$m, and the cell length is 10 cm in the inner region or 25 cm in the outer region. The single point resolution in $\phi$ is roughly pitch/ $\sqrt{12}$ assuming a uniform distribution of the hits. The strip tracker has 9.3 million strips in total corresponding to 198 $m^2$ of active silicon area. The silicon strip tracker consists of three subsystems as shown in Fig. 3.3: the Tracker Inner Barrel and Disks (TIB/TID), the Tracker Outer Barrel (TOB), and the Tracker End-Caps (TEC).

The TIB/TID surrounds the pixel detector and occupies the radial region between 20 cm and 55 cm. The TIB has 4 concentric barrel layers extending in $z$ between $\pm$70 cm. The TID consists of 3 disks at each end positioned between $z = \pm 80$ cm and $z = \pm 90$ cm, covering the radii between 20 cm and 55 cm. The sensors used in TIB/TID are 320 $\mu$m thick. The strip pitch is 80 $\mu$m on the innermost two layers of TIB and 120 $\mu$m on TIB layer 3 and 4. The pitch varies between 100 $\mu$m and 141 $\mu$m in the TID.

The TOB is outside of TIB/TID and has an outer radius of 116 cm. It has 6 barrel layers covering $|z| < 118$ cm. In this region with larger radii, the strip pitch is increased to 183 $\mu$m for the first 4 TOB layers and 122 $\mu$m for the outer 2 layers. In order to cover a larger area and at the same time limit the number of readout channels, the length of the strip sensors has to be increased as well. The increased strip length leads to a higher electronics noise due to increased strip capacitance. Therefore, thicker silicon sensors (500 $\mu$m) are used for this outer region in order to reduce the capacitance and maintain a high signal-to-noise

ratio.

The TECs are composed of 9 disks on each side occupying the region between 124 cm $< |z| <$ 82 cm and 22.5 cm $< r <$ 113.5 cm. Each disk contains up to 7 rings of strip sensors, with their pitches varying between 97 $\mu$m and 184 $\mu$m. The sensor thickness on the inner 4 rings is 320 $\mu$m and is 500 $\mu$m on the outer 3 rings.

The two innermost layers and rings of TIB/TID and TOB, together with ring 1, 2, 5 of the TECs, use stereo modules that combine two layers of strip sensors. Two micro-strip modules are mounted back-to-back with the second module rotated by a small stereo angle of 100 mrad in order to achieve a better measurement of $z$ coordinate in barrel or $r$ coordinate in disks. The back-to-back modules are indicated in Fig. 3.3 as double lines. This tracker layout ensures at least 9 hits in most of the tracker acceptance region ($|\eta| < 2.4$), of which at least 4 are stereo hits from the back-to-back modules.

The strip tracker adopts an analog readout scheme. Signals from the silicon sensors are amplified, processed, and buffered by a custom ASIC named APV25. Four or six APV25 readout chips are mounted onto a multi-chip module (hybrid), which is in charge of distributing power, routing clock, control, and data lines, as well as transferring heat into the cooling system. Should a trigger signal (c.f. Section 3.2.7) be received , the APV25 chip transmits the analog signals via optical links to the Front End Driver (FED) boards outside of the experiment cavern. Each FED receives data from 96 optical links and then digitize them using a 40 MHz, 10 bit ADC. Pedestal corrections and common mode subtractions are applied before the output list of clusters with address information and signal height is sent to the central data acquisition system (DAQ).

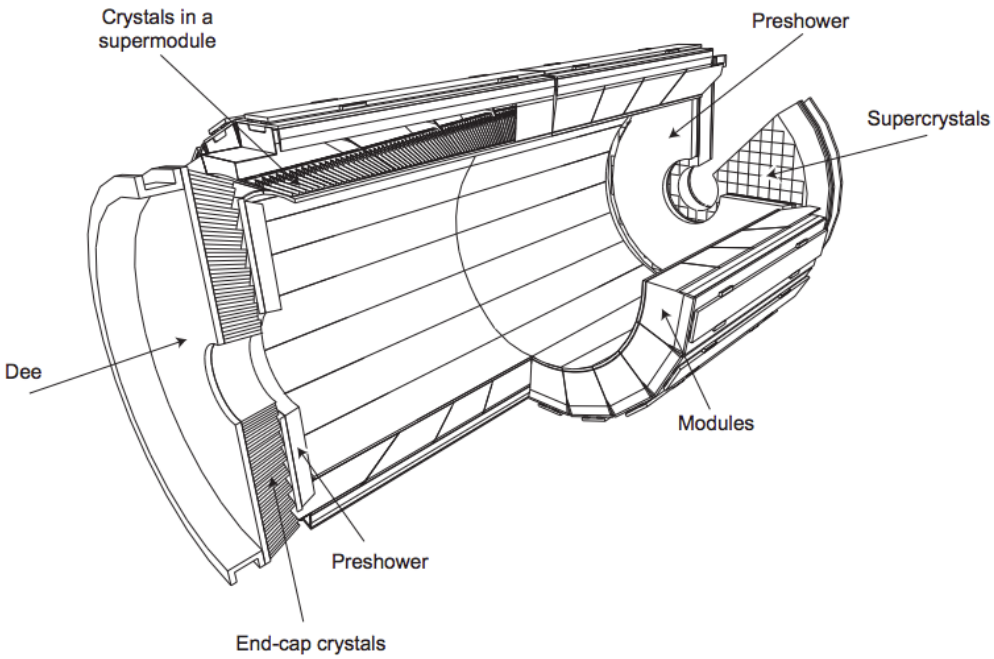


Figure 3.4: Layout of the CMS electromagnetic calorimeter [32].

### 3.2.4 Electromagnetic calorimeter

The CMS electromagnetic calorimeter (ECAL) [36] is a hermetic homogeneous calorimeter composed of lead tungstate ($PbWO_4$) crystals. It is designed to stop electrons and photons produced from the proton collisions and measure their energy. The ECAL consists of a barrel part (EB) covering $|\eta| < 1.479$ and endcaps (EE) on each side covering the range $1.479 < |\eta| < 3.0$. In addition, a preshower detector is placed in front of each EE covering $1.653 < |\eta| < 2.6$ in the endcap region. The layout of the ECAL is shown in Fig. 3.4.

The $PbWO_4$ crystal used in the ECAL has a high density of 8.28 $g/cm^3$, with a radiation length of 0.89 cm and a Moliére radius of 2.2 cm. This allows the design of a compact calorimeter with fine granularity. The crystal is optically clear, radiation hard, and has a short scintillation decay time close to the LHC

bunch crossing time. A total of 61200 $PbWO_4$ crystals are installed in the EB, and 7324 crystals are used in each EE. The crystal is taper shaped, $22\times22$ mm$^2$ ($28.6\times28.6$ mm$^2$) at the front face and $26\times26$ mm$^2$ ($30\times30$ mm$^2$) at the rear face with a length of 23.0 cm (22.0 cm) in the EB (EE). They are mounted pointing roughly to the interaction point. In the EB, crystals are assembled into supermodules, each of which contains 1700 crystals. Crystals in the EE are grouped into $5 \times 5$ units named supercrystals (SCs).

The scintillation light from electromagnetic showers in the crystals is detected by avalanche photodiodes (APD) in the EB and vacuum phototriodes (VPT) in the EEs. Different choices of photodetectors are due to different configurations of the magnetic field as well as the expected radiation level in the barrel and endcap regions. In the EB, each crystal is read out by a pair of APDs. Each APD has an active area of $5 \times 5$ mm$^2$ and is operated at a gain of 50. The VPTs used in the EEs are developed specially to operate in the CMS 3.8 T magnetic field. A VPT has a diameter of 25 mm and an active area of about 280 mm$^2$. Each VPT is glued to the back of one crystal.

A trigger tower consists of a set of $5\times5$ crystals in the EB or EE. Each trigger tower is read out by the on-detector electronics, where the signals are shaped and digitized. Trigger primitives are created in the off-detector electronics based on the trigger towers. Each trigger primitive consists of the sum of transverse energy ($E_T$) deposited in the trigger tower as well as the shower shape information. A selective readout scheme is used in the ECAL to reduce the large data volume. Each trigger tower is categorized based on its deposited $E_T$ into high ($E_T > 5$ GeV), medium ($E_T > 2.5$ GeV), and low interest classes . A trigger tower is read out with no zero suppression if it is classified as at least medium

interest class. In case the trigger tower is of high interest class, all its neighboring trigger towers are also read out with no zero suppression. In other cases, the crystals are read out with zero suppression at about three standard deviation of the noise level.

The preshower detector (ES) is a 20 cm thick sampling calorimeter placed in front of the EE on each side, as shown in Fig. 3.4. It is composed of four alternating layers of lead radiators initiating electromagnetic showers and silicon strip sensors to measure the energy and profiles of the shower. The strips used in the ES have a 1.9 mm pitch, and the two strip layers are oriented in orthogonal directions. The ES is designed to provide better neutral pion rejections in the barrel-endcap transition regions. Electron identification against minimum ionizing particles and position measurement of electrons/photons also benefit from the preshower detector.

The ECAL relative resolution can be parameterized as a function of energy as [32]

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2 \tag{3.3}$$

The first term is the stochastic term. Event-to-event fluctuations in the lateral shower containment, photostatistics, and fluctuations in preshower energy measurements are three main contributions to the stochastic term. The second term is the noise term primarily due to electronics noise, digitization noise, and pileup noise. The last constant term represents contributions from non-uniformity of the longitudinal light collection, errors in inter-calibration, and energy that leaks from the back of the ECAL. The above equation is valid for energies up to about 500 GeV, beyond which energy leakage at the rear of ECAL becomes significant.

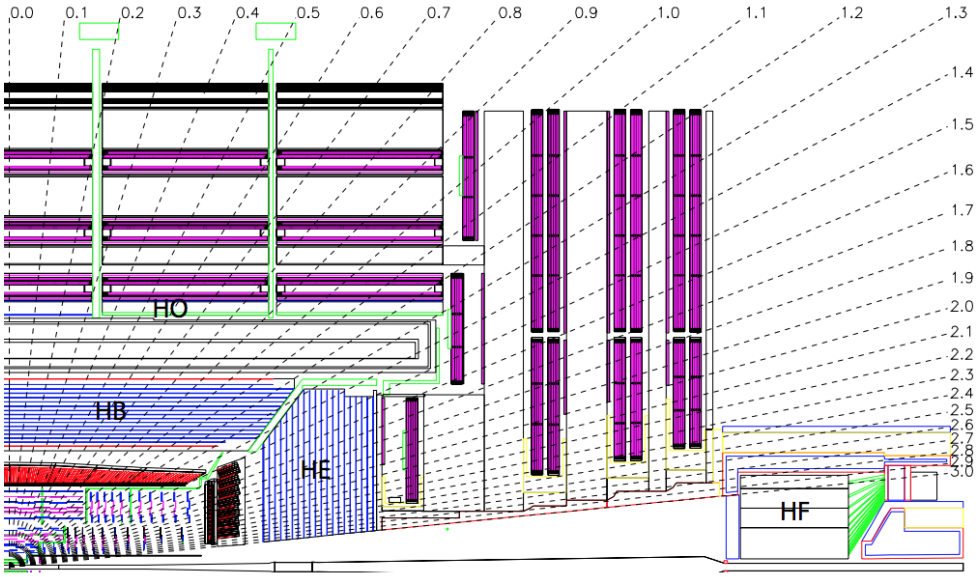


Figure 3.5: Schematic of a quarter of the CMS detector in *r*-*z* view [32]. The HCAL subsystems are labeled in the diagram.

### 3.2.5 Hadron calorimeter

The CMS hadron calorimeter (HCAL) [37] is a sampling calorimeter made from alternating layers of brass/steel absorbers and plastic scintillators. It measures the energy of hadrons and also plays an important role in determining missing transverse energy due to neutrinos or exotic particles. The HCAL sits outside of the ECAL and includes four components: hadron barrel (HB), hadron endcap (HE), outer calorimeter (HO), and forward calorimeter (HF), as shown in Fig. 3.5.

The HB is placed between the EB and the CMS solenoid covering $|\eta| < 1.3$. It is segmented into 18 $\phi$ wedges, each if which contains a total of 16 absorber layers. The absorber layers are made of 50.5 mm or 56.5 mm brass plates with the exception that the innermost and the outermost absorber layers are made of stainless steel for structural strength. Each wedge is further divided into 4

$\phi$ and 16 $\eta$ segments, with a granularity of (0.087, 0.087) in $(\eta, \phi)$. There are 16 layers of scintillators. The innermost scintillator layer is made of 9-mm-thick Bicron BC408 and is placed in front of the first layer of the absorber in order to sample the hadronic showers developing in the inert material outside of the EB. The other layers use Kuraray SCSN81 plastic scintillators with a thickness of 3.7 mm except for the outermost layer, in which a larger thickness of 9 mm is used to account for the late developing showers. The scintillation tiles in each $\phi$ segmentation of a given layer are grouped into one tray for readout. Green double-cladded wavelength shifting fibers (WLS) placed in machined grooves of the tiles are used to collect scintillation light. Signals are sent out via the clear fibers that are spliced to the end of WLS to a hybrid photodiode (HPD), where optical signals are converted to electrical signals. The analog signals from the HPD are then digitized in the front-end electronics before they are transmitted off the detector.

The HE is positioned between the EE and the muon endcap iron yoke covering the pseudorapidity range $1.3 < |\eta| < 3.0$. The HE has 17 brass absorber plates with a thickness of 79 mm and an outermost 10-cm-thick support plate made of stainless steel. Similar to the HB, Bicron BC408 with a thickness of 9 mm is used in the innermost scintillator layer, and 3.7-mm-thick SCSN81 plastic is used for the other 17 layers. The HE is divided into 14 $\eta$ sectors including one shared $\eta$ sector with the HB. The granularity in $(\eta, \phi)$ of the HE is (0.087, 0.087) for $|\eta| < 1.6$ and roughly (0.17, 0.17) for $|\eta| > 1.6$.

The radial size of the HB, thus the total amount of absorber material, is limited by the size of the CMS solenoid. As a result, an outer calorimeter is installed outside of the solenoid as a tail catcher to measure the shower energy after the

HB. The HO uses the solenoid coil as an additional absorber and consists of a single scintillator layer made of 10-mm-thick Bicron BC408 attached to the outer surface of the first layer of the iron yoke at radius $r = 4.07$ m. In the central region between $z = +1.268$ m and $z = -1.268$ m, an additional scintillator layer is add to the inner surface of the iron yoke at radius $r = 3.82$ m to compensate for the minimal absorber depth of the HB around $\eta = 0$.

The acceptance of the HCAL is extended by the HF in the forward region to $|\eta| < 5$. The HF is located 11.2 m away from the the interaction point. The cylindrical absorber structure has a depth of 165 cm with an inner radius of 12.5 cm and an outer radius of 130.0 cm. Radiation hard quartz fibers are used as scintillators and are inserted in the 5-mm-thick grooved absorber plates made of steel. Only half of the fibers are embedded in the full depth of the absorber, the other half does not cover the first 22 cm of the absorber. In this way, electromagnetic showers, which deposit most of their energy in the first 22 cm therefore produce higher signal in the longer fibers, are distinguished from the hadron showers. Light signals in fibers of an $\eta$-$\phi$ tower are combined and sent to a photomultiplier tube (PMT). The HF detector and its front-end readout electronics are protected by a hermetic radiation shielding composed of steel, concrete, and polyethylene.

Several components of the HCAL described above have been upgraded [38] since LHC Run 1. The performance of HPDs in HB, HE, and HO was degraded over the LHC Run 1 due to electrical discharge and gain variation in pixels. The HPDs in the HO, where the effect was more severe, were replaced with silicon photomultipliers (SiPMs) during Long Shutdown 1 (LS1) in 2013 and 2014. The same replacement is planned for HB and HE during Long Shutdown 2 (LS2)

after the LHC Run 2. As for the HF, the original PMTs were replaced during the extended year-end technical stop between 2016 and 2017 by multi-anode PMTs with thinner optical windows that can provide better identifications of spurious signals. The front-end electronics was also upgraded during this period, enabling high precision timing measurements.

### 3.2.6 Muon system

The CMS muon system [39,40] is the outermost subsystem positioned outside of the CMS superconducting solenoid. It provides precise measurements of muon tracks as muons travel through the magnetic field. The muon system is composed of three types of gaseous detectors: drift tubes (DT), cathode strip chambers (CSC), and resistive plate chambers (RPC). A quarter of the muon system in $r$-$z$ view is in Fig. 3.6.

#### Drift Tubes

The relatively low muon rate, small neutron background, and uniform magnetic field in the barrel region allow the use of drift tube chambers with standard rectangular cells. The DT covers $|\eta| < 1.2$ in the barrel region. It is segmented into 12 $\phi$ sectors, each of which consists of 4 stations (MB1, MB2, MB3, MB4 in Fig. 3.6) with each station sandwiched between the iron yoke plates. The parallel drift cells are grouped into layers, and 4 staggered layers are further grouped into a superlayer (SL). Each DT chamber, as shown in Fig. 3.7 on the left, consists of two SLs along the $z$ direction to measure $r$-$\phi$ coordinates and one SL (except in MB4) orthogonal to the other two SLs in order to provide a

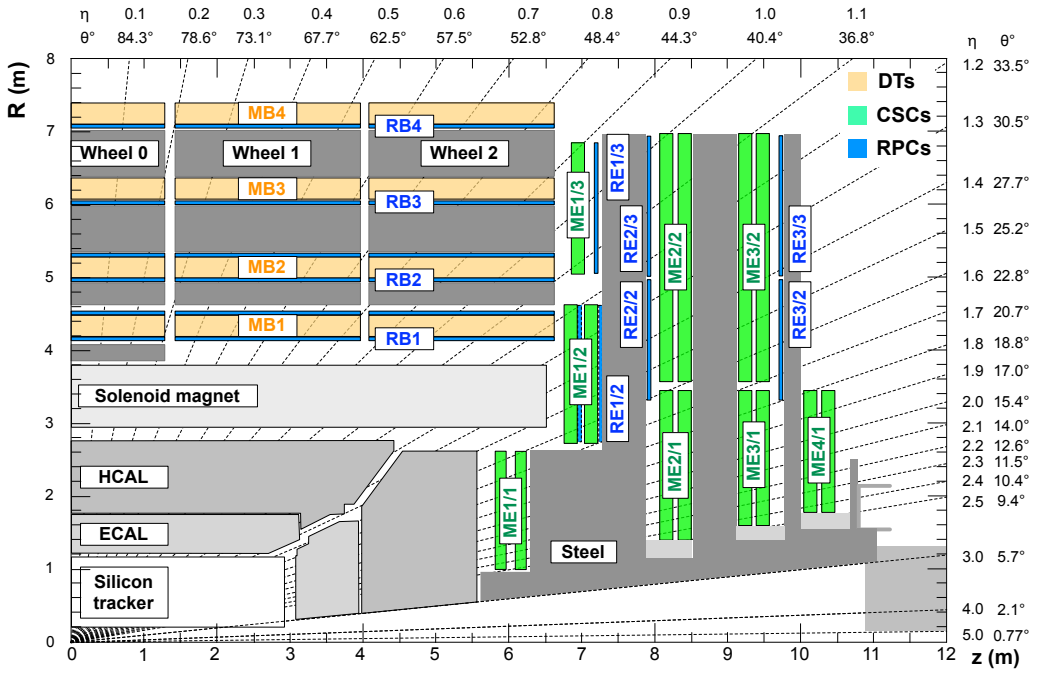


Figure 3.6: Cross section of a quarter of the CMS muon system in *r*-*z* view [40].

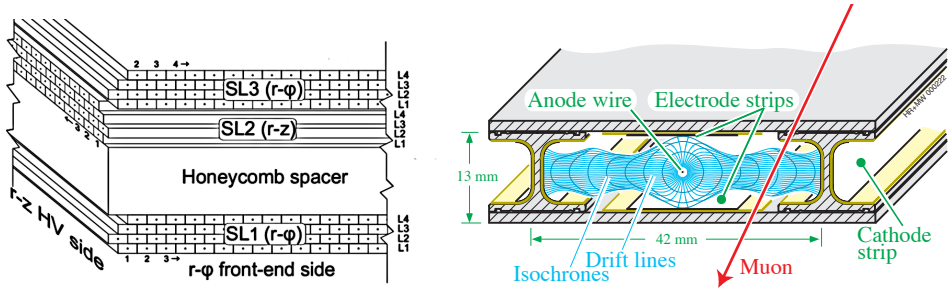


Figure 3.7: Left: the schematic of a DT chamber showing orthogonal orientations of the drift cells in different layers. Right: a DT cell with the drift lines [40].

measurement in $z$.

A sketch of the drift cell is shown in Fig. 3.7 on the right. Each drift cell has a $42 \times 13$ mm$^2$ rectangular cross section. A high voltage of +3600 V is applied to the anode at the center made of a gold-plated stainless-steel wire with a diameter of 50 $\mu$m. Two electrode strips above and below the anode wire set to +1800 V and two cathodes on the side walls set to -1200 V are used to shape the effective drift field. The drift cells are filled with a gas mixture of

85% Ar and 15% $CO_2$. The saturated drift velocity is about 55 $\mu$m/ns, giving a maximum drift time of roughly 380 ns.

**Cathode Strip Chambers**

Due to higher muon rate, larger backgrounds, and non-uniform magnetic field in the endcap regions, cathode strip chambers are used because they have fast response time and can be finely segmented. The CSC on each side contains 4 stations sandwiched between the endcap flux return yoke disks covering the range of $0.9 < |\eta| < 2.4$). The CSCs are multi-wire proportional chambers perpendicular to the beam pipe made of alternating layers of trapezoidal cathode panels and anode wire planes, shown in Fig. 3.8. There are seven cathode panels per CSC chamber, each of which contains 80 cathode strips along the radial direction providing precise $r$-$\phi$ coordinate measurements. Six gas gaps are formed from he cathode panels, containing anode wires perpendicular to the cathode strips enabling a coarse measurement in $r$. A gas mixture of 50% $CO_2$, 40% Ar, and 10% $CF_4$ is used in all CSCs. A CSC has a length in radial direction ranging between 1.7 m to 3.4 m and covers a $\phi$ angle of about $10^\circ$ or $20^\circ$.

**Resistive Plate Chambers**

The resistive plate chambers are added as a complementary trigger system of the CMS muon system with a goal of providing fast, independent, and highly-segmented trigger. The RPCs provide timing measurements with excellent resolution of about 1 ns but relatively coarse position measurements compared to the DT and CSCs. They are installed in both barrel and endcap regions, as

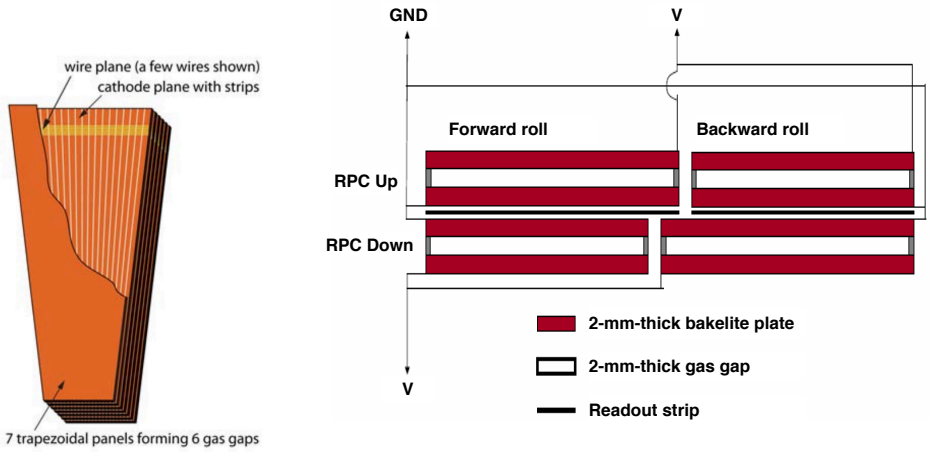


Figure 3.8: Left: cut-away view of a CSC with 7 trapezoidal cathode strip panels and 6 gas gaps with anode wires [32]. Right: schematic of a RPC in barrel [40].

shown in Fig. 3.6. Six layers of RPCs are in the barrel region attached to the DT chambers, and three layers are embedded in the endcap on each side.

The RPCs are made of double-gap chambers ("RPC Up" and "RPC Down"). The layout of a PRC is shown in Fig. 3.8. Each gap is 2 mm thick and is filled with gas mixture of 95.2% $C_2H_2F_4$, 4.5% i-$C_4H_{10}$, and 0.3% $SF_6$. Each gap is enclosed by two parallel 2-mm-thick resistive Bakelite plates. A voltage of 9600 V is applied to the top (bottom) plates of RPC Up (Down). The readout strips are placed between the upper and lower gaps to sample the sum of the signals in the two gaps. The signals are further discriminated and shaped by the front-end electronics attached to the end of the strips.

### 3.2.7 Trigger and data acquisition system

The LHC collides proton bunches at a rate of 40 MHz with an average pileup of about 20 during Run 1 and about 30 during Run 2. An enormous amount of

data is produced, in which, however, only a very small fraction is of primary interest to the CMS physics program (e.g. productions of Higgs bosons, weak gauge bosons, top quarks, or new particles beyond the standard model). It is also impractical for the CMS experiment to store and process all collision data due to the limited readout bandwidth and the finite computing power. Therefore, a trigger system is used in the CMS experiment to achieve a large data reduction and at the same time select the events that are most likely to come from interesting physics for further studies. CMS adopts a two-level trigger system consisting of a Level 1 (L1) Trigger and a High-Level Trigger (HLT). The L1 Trigger is based on customized hardware and reduces the data rate from 40 MHz to about 100 kHz. The HLT is a software trigger running on a farm of processors and further reduces the rate to around 1 kHz before data storage.

**Level 1 Trigger**

As the first level of the trigger system, the L1 Trigger [41] makes a decision to either accept or reject a collision event within 3.2 $\mu$s using information with reduced precision from the calorimeters (ECAL, HCAL, and HF) and the muon system (DT, CSC, and RPC).[3] The local trigger primitives provided by the subdetectors are combined in several stages, before a global Level 1 Accept (L1A) decision is made and sent to the subdetectors. An overview of the L1 trigger systems is shown in Fig. 3.9.

The calorimeter trigger primitives are generated based on the trigger towers. A trigger tower is $5 \times 5$ crystals in ECAL or HCAL readout tower with an $\eta$-$\phi$ coverage of $0.087 \times 0.087$. The sum of the transverse energy ($E_{\mathrm{T}}$) in each trig-

[3]The silicon tracker is currently not used in the L1 Trigger. A proposal to include the tracker in the L1 Trigger for the CMS Phase 2 upgrade is discussed in details in Chapter 6.

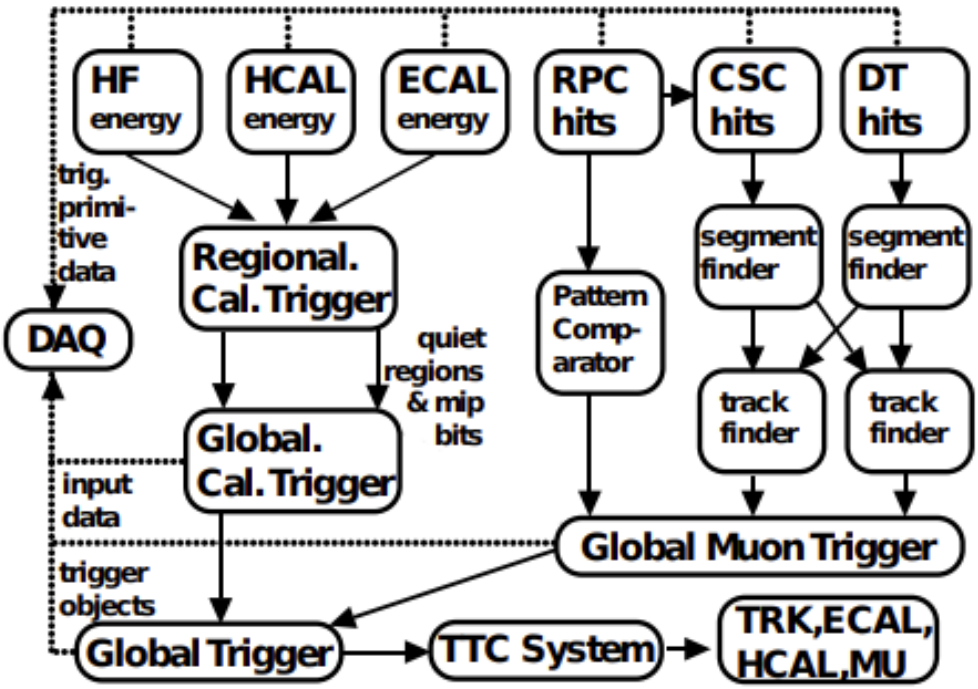


Figure 3.9: Data-flow for the Level 1 Trigger [41].

ger tower is calculated in the Trigger Primitive Generators (TPG) integrated in the calorimeter front-end electronics. The calorimeter trigger primitives from ECAL, HCAL, and HF in the same $\eta$-$\phi$ regions are grouped and transmitted to the corresponding Regional Calorimeter Triggers (RCT). The RCT determines regional trigger information including electron/photon candidates, sum of $E_{\mathrm{T}}$, isolation, etc. The regional calorimeter trigger objects are forwarded to the Global Calorimeter Trigger (GCT), where the jets, the total $E_{\mathrm{T}}$, the scalar sum of all jets (above some programmable threshold) transverse energy ($H_{\mathrm{T}}$), and ranked electron/photon candidates are produced.

The local components of the muon trigger include track segments and hit patterns from the DT as well as 3D track segments from the CSCs. These track segments from the DT and CSCs are joint in the Track Finders, where muon

tracks are reconstructed. As a dedicated muon trigger system with very good time resolution, the RPCs also provide track candidates to the muon trigger. The Pattern Comparator Trigger (PACT) compares the 4D (space and time coordinates) PRC hits with predefined patterns to find tracks. The track candidates from all three muon systems are combined in the Global Muon Trigger (GMT). By using the complementary and redundant information from the three subsystems, GMT achieves improved trigger efficiency and reduced background rates.

Finally, the trigger objects from the GCT and GMT are sent to the Global Trigger (GT). A total of 128 trigger algorithms can be executed in parallel using the input trigger objects in the GT. Details on the trigger algorithms and their performance can be found in Ref. [42]. A L1A decision is made based on the deployed trigger algorithms as well as the readiness of the subdetectors and the DAQ. The L1A signal is communicated back to the subdetector front ends via the Timing, Trigger, and Control (TTC) system for readout.

**Data Acquisition and High-Level Trigger**

When the detector front ends receive the L1A signals, the data stored in the pipelined buffers of the on-detector electronics are transferred to the FEDs at the maximum rate of 100 kHz. The data in the FEDs are read out and merged by the Front-end Read-out Links (FRL). The event fragments from the FRLs are assembled and sent out by the Event Builder to the filter systems in the surface building. Each event fragment has a size of 2 kB on average. With about 500 FRLs, the data throughput from the Event Builder to the filter systems is about 100 GB/s.

The High-Level Trigger algorithms are executed in the filter systems to further reduce the event rate to about 1 kHz and select events from interesting physics to be stored for offline processing. The event selections at the HLT are based on a similar but faster version of the offline reconstruction software. The filter farm consists of $O(10000)$ CPU cores and can process an event within 200 ms. More details on the DAQ and the HLT including its algorithms and performance can be found in Ref. [42, 43].

CHAPTER 4

# OBJECT RECONSTRUCTION

This chapter describes the reconstruction and identification of the physics objects at CMS that are relevant to the $t\bar{t}H$ analysis. Section 4.1 describes the particle-flow algorithm for global event reconstruction used by CMS. The reconstruction of charged particle trajectories in the silicon tracker is discussed in Section 4.2, followed by a brief description of reconstructing the collision vertices and beam spots in Section 4.3. Electron and muon reconstructions as well as the basic identifications are discussed in Section 4.4 and Section 4.5 respectively. Section 4.6 presents the jet clustering and reconstruction algorithms based on PF objects. Algorithms for tagging jets originating from b quarks are discussed in Section 4.7. The reconstruction and identification of hadronically decaying tau leptons are described in Section 4.8. Finally, the missing transverse momentum is discussed briefly in Section 4.9.

## 4.1 Particle-flow reconstruction

The various information from all CMS subdetectors, described in Section 3.2, are combined using a particle-flow (PF) algorithm [44] to reconstruct and identify all individual particles in an event. The fine granularity of the CMS subdetectors makes it possible to use the PF algorithm for reconstructing events from proton-proton collisions. The building blocks of the PF algorithm, namely PF elements, include charged particle tracks and calorimeter clusters. The reconstruction of tracks in the silicon tracker is described in Section 4.2. Specialized track reconstructions for muons and electrons are discussed in Section 4.4 and 4.5 respec-

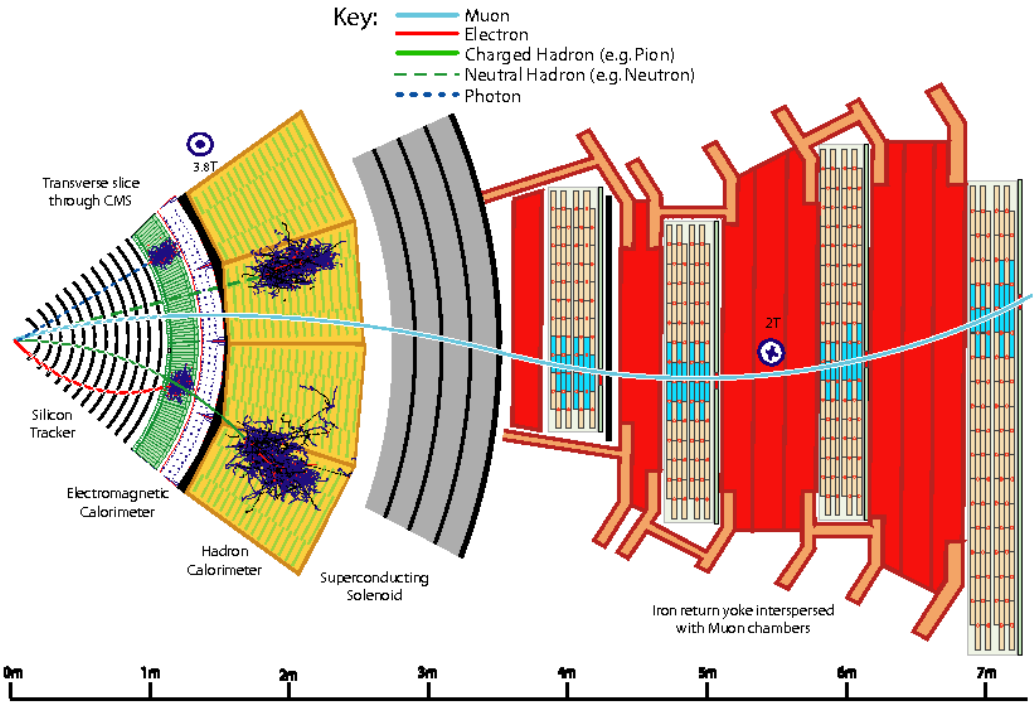


Figure 4.1: A transverse view of a slice of the CMS detector showing the interactions of different types of particles identified by the particle-flow algorithm with subdetectors [44].

tively. Clustering of energy deposits for the PF algorithm is performed independently in the barrel and endcaps of ECAL and HCAL. The cluster seeds are selected as the calorimeter cells with energy deposits larger than that of their neighbors and above a given threshold. The PF clusters are then formed around the seeds by iteratively adding adjacent cells if the energy deposits in the cells are at least twice of the noise level. Different PF elements are linked into PF blocks based on their geometrical and kinematic compatibility. Details of the link algorithm can be found in Section 4.1 of Ref. [44]. In each PF block, particle candidates are reconstructed using PF elements. The PF elements that correspond to a successfully reconstructed particle are masked in the PF block and are not used in reconstructing other particles.

Different types of particles, namely muons, electrons, photons, charged and neutral hadrons, are identified by the PF algorithm. Fig. 4.1 shows a sketch in the transverse plane of different types of particles interacting with different CMS subdetectors. Muons are identified by connecting tracks reconstructed in the muon system and the inner tracker. Electrons are identified as tracks and the associated ECAL clusters that match to the track extrapolations to the ECAL and to the bremsstrahlung photons emitted along their trajectories. In case there are ECAL clusters that do not match any tracks in the silicon tracker, the ECAL clusters are identified as photons. Charged hadrons are identified as tracks that are linked to both ECAL and HCAL clusters. Neutral hadrons are identified as HCAL and ECAL clusters that are not linked to any tracks. In case a neutral hadron overlaps with charged particles, the neutral hadron is detected as a calorimeter energy excess compared to the expected total energy deposits from the associated charged particle. The PF algorithm does not distinguish the species of the charged or neutral hadrons.

## 4.2 Tracks

The charged particles produced in proton-proton interactions follow approximately helical trajectories in the axial magnetic field generated by the CMS solenoid. The CMS silicon tracker, described in Section 3.2.3, detects the hits on the trajectories when the charged particles traverse layers of the silicon pixel and strip sensors. With an instantaneous luminosity of $O(10^{34})\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ at the LHC, $O(1000)$ tracks are expected per bunch crossing in the tracker volume. Dedicated software algorithms are developed for reconstructing tracks in this challenging environment to fully exploit the capability of the CMS tracker in or-

der to achieve excellent tracking performance [45]. Figure 4.2 shows the relative $p_T$ resolution as a function of $p_T$ and the transverse impact parameter $d_0$ resolution as a function of $\eta$ for tracks in simulated $t\bar{t}$ events with an average pileup of 35 in 2016 and 2017 detector conditions at 13 TeV.

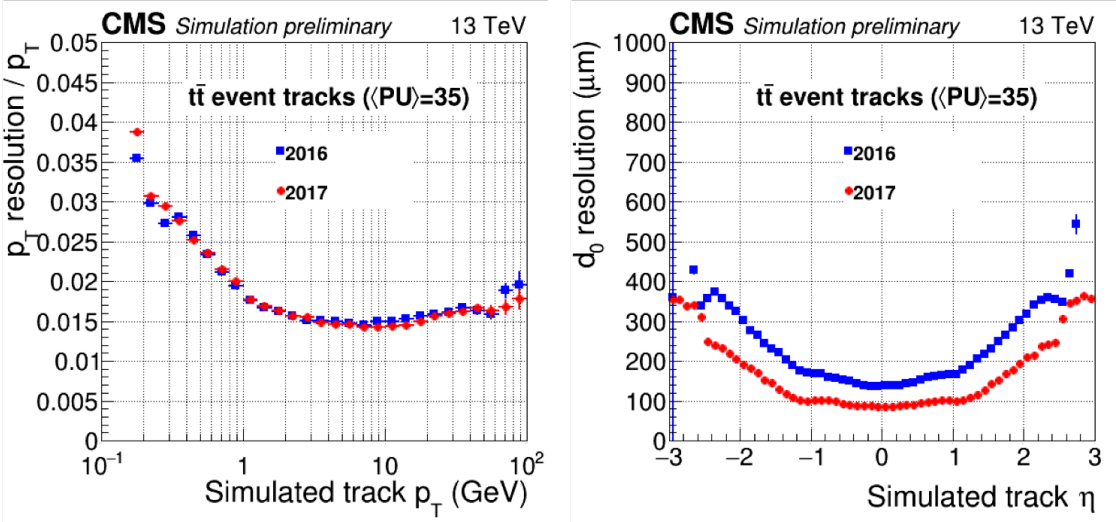


Figure 4.2: Left: resolution of the relative $p_T$ as a function of simulated track $p_T$. Right: resolution of the transverse impact parameter $d_0$ as a function of simulated track $\eta$. Both resolution plots are based on simulated $t\bar{t}$ events with an average pileup of 35. Performances in both 2016 (blue) and 2017 (red) detector conditions are shown [46].

The track reconstruction process starts with the reconstruction of local hits in the pixel and strip sensors. For pixels, the signal in each pixel is read out if the deposited charge is above a certain threshold.Hit clusters are formed offline by grouping adjacent pixels. Two algorithms are used for determining the coordinates, as well as the associated uncertainties, of the pixel clusters. A faster algorithm uses the geometrical center of the cluster corrected by the differences of charge deposits in the pixels at the edge of the cluster. This faster algorithm is used in the seeding and pattern recognition stages of the track reconstruction. A more precise algorithm used in the final track fit stage determines the cluster position by comparing the charge distribution in the cluster to the templates

generated in the simulation with various particle incident angles and irradiated pixel sensors. In the strip detector, a cluster is formed by grouping neighboring strips passing certain signal-to-noise ratio thresholds. The position of the cluster is determined based on the charge-weighted average of the strip positions.

The reconstructed hits are then used to estimated the momentum and position parameters of the tracks. The tracking software used by CMS is named Combinatorial Track Finder (CTF). The CTF algorithm is based on the Kalman filter (KF) method [47–50]. Multiple iterations of the CTF sequence are run to produce a collection of tracks. Each iteration targets at a specific type of tracks, starting from the easiest ones (e.g. high $p_{\mathrm{T}}$, prompt tracks) to more challenging collections (e.g. low $p_{\mathrm{T}}$, displaced tracks). If the hits are associated with the reconstructed tracks, they are removed for the next iterations. In this way, the complexity of reconstructing more difficult tracks in the subsequent iterations are reduced.

Each iteration of the CTF consists of four steps: seed generation, track finding, track fit, and track selection. Track seeds are formed with either three 3D hits or two 2D hits with beam spot constraint. A 3D hit is either a pixel hit or a stereo hit in the strip back-to-back modules (cf. Section 3.2.3). Only hits in the inner parts of the track are used as seeds, and the track candidates are built outwards. Initial estimates of the five helix parameters, as well as their uncertainties, are calculated from the seed.

The track finding step extrapolates the trajectory estimated from the seed and searches for associated hits in other layers. Multiple scattering and energy loss are taken into account for the extrapolation. Hits that are compatible with the projection are added to build track candidates and are used to update

the track parameters. Up to 5 track candidates are retained at each layer for a given seed in order to keep a manageable number of track candidates. The track candidates are projected to the next layers, and the extrapolation process is repeated until it reaches the end of the tracker or if the track candidates fail some kinematic or quality requirements. At the end of the track finding step of each iteration, duplicate track candidates are removed based on the number of shared hits.

A final fit is performed using all the hits of a track candidate in order to take advantage of the full information of the trajectory and to remove potential bias from beam spot constraints. The fit uses a Kalman filter and smoother. The innermost hits are used to initialize the Kalman filter, and the fit updates the track parameters iteratively by adding the hits from the inside outwards. A second filter, initialized with the results from the first filter, smooths the track parameters by fitting the hits again from the outside inwards toward the beam line. The track parameters at a given layer is obtained from the weighted average of the two filters. The extrapolation in the fitting step uses a Runge-Kutta propagator, which takes both the material effect and inhomogeneous magnetic field into consideration. After the fit, spurious hits are removed if they are incompatible with the track based on $\chi^2$. The track is then fitted again with the outliers removed. The entire process is repeated until there are no outliers or the track is discarded due to insufficient number of good hits.

Finally, a selection of the tracks is applied based on the number of hits, $\chi^2$/dof from the fit, and the significance of impact parameters. The selection step reduces significantly the fraction of the reconstructed tracks that are not associated with any charged particle.

## 4.3 Vertices and beam spot

The proton-proton interaction vertices can be reconstructed by clustering the tracks based on the $z$ position of their closet approach to the beam line. A deterministic annealing algorithm is used for clustering tracks, and the positions of the candidate vertices are determined from an adaptive vertex filter [45]. The primary vertex (PV) of an event is defined as the hard-scattering vertex, while others are pileup vertices. The PV is identified as the vertex with the largest $\sum p_{\mathrm{T}}^2$ of its associated tracks during LHC Run 1. Due to the increased instantaneous luminosity in Run 2, an improved version of PV identification has been developed. The tracks associated with a vertex are clustered into jets using the anti-$k_t$ jet finding algorithm [51] with a cone size parameter of 0.4. The $\sum p_{\mathrm{T}}^2$ of each vertex is then computed using the jets, unclustered tracks, and the associated missing transverse energy, which is calculated as the negative vectorial sum of the jet transverse momenta. The vertex with the largest $\sum p_{\mathrm{T}}^2$ is identified as the PV.

The beam spot is the region where the proton beams collide. The center position and the spread of the beam spot are determined by fitting a collection of the reconstructed collision vertices in every luminosity section (LS). A LS corresponds to collision events recorded during 23 seconds of LHC operation.

## 4.4 Muons

Muons within the pseudorapidity range $|\eta| < 2.4$ are detected by the CMS muon system, described in Section 3.2.6. All other particles except neutrinos are ex-

pected to be absorbed by the upstream calorimeters and iron yoke plates in the muon system. The muon tracks are first reconstructed separately in the silicon tracker (tracker tracks) and in the muon system (standalone muon track), before they are combined to reconstruct muon objects. Tracker tracks are built as described in Section 4.2. The reconstruction of standalone muon tracks uses information from all three muon subdetectors. Track segments are formed by clustering hits within the DT or CSC and are used to seed the pattern recognition step. Matched hits and segments in DTs, CSCs, and RPCs are then fitted using a Kalman filter to build the standalone muon tracks.

The muon tracks can be built by either an "inside-out" or an "outside-in" approach. In the "inside-out" reconstruction, tracker tracks are propagated to the DT or CSCs. Only tracker tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ and total momentum $p > 2.5\,\mathrm{GeV}$ are used in the extrapolation. The extrapolated positions are matched to the local track segments in the muon system based on the distances in the $r$-$\phi$ plane. If a tracker track matches to at least one muon segment, a "tracker muon" is built. The "outside-in" approach matches each standalone muon track to tacks built in the silicon tracker. The track parameters of both tracks are propagated to a common surface, and a match is found if the track parameters are compatible. The hits of the matched standalone muon track and the inner track are then combined and fitted using a Kalman filter to build a "global muon".

A set of criteria based on particle-flow are applied to identify muons. Two muon identifications, loose and medium muon ID [52], are relevant in the $\mathrm{t\bar{t}H}$ analysis presented in this thesis. The loose muon ID requires that muons are identified by the PF algorithm and are either tracker or global muons. It aims to

distinguish genuine muons, either prompt ones produced at the primary vertex or the non-prompt ones from hadron in-flight decay, against charged hadrons that are not stopped in the calorimeters and are misidentified as muons. The medium muon ID adds additional requirements on top of the loose ID in order to only identify prompt muons or muons from heavy flavor decay. The efficiency of medium muon ID is shown in Fig. 4.3 for both data and simulation in 2017. The identification variables include the quality of the global track fit, the multiplicities of hits per track, the compatibility of the tracker tracks with the standalone muon tracks, muon segment compatibility, and $\chi^2$ returned by a kink-finding algorithm. The muon segment compatibility quantifies the agreement of the tracker track extrapolation to the muon system with the muon segments based on the number and consistency of the matches. The kink-finding algorithm checks if a tracker track is compatible with being a single track by systematically splitting the track into two and comparing the two separate tracks.

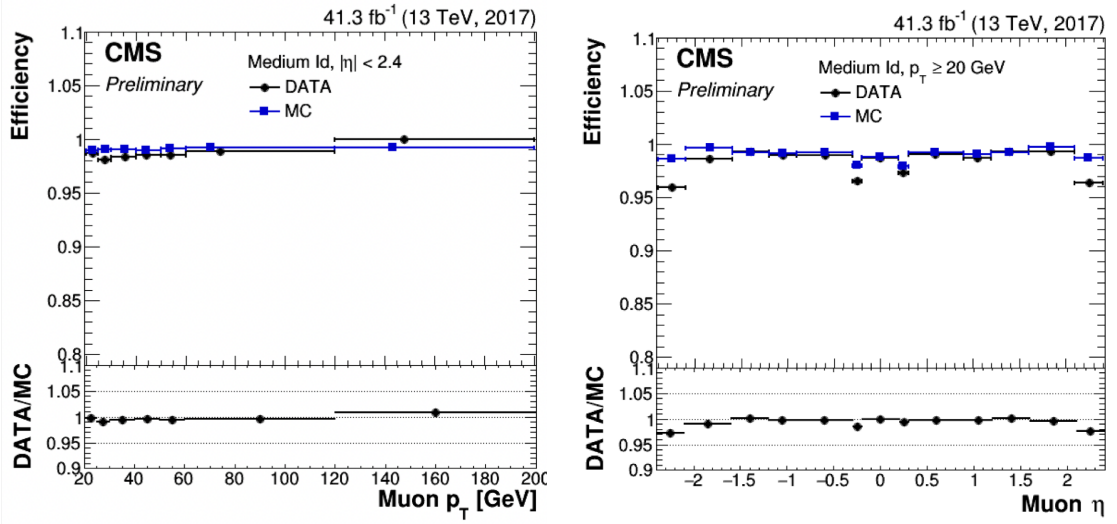


Figure 4.3: Muon medium ID efficiency as a function of muon $p_{\mathrm{T}}$ (left) and $\eta$ (right) for 2017 data (black) and simulation (blue) [53]. The denominator for the efficiency calculation is all tracker tracks with $p_{\mathrm{T}} > 20$ GeV.

Additional muon isolation and identification requirements are applied in the context of the $t\bar{t}H$ analysis presented in this thesis in order to further separate prompt muons from the non-prompt ones produced in decays of charm and bottom quarks. Because this set of selections is closely coupled to the background estimation strategy of the analysis, details of the dedicated prompt lepton selection are given in Section 5.2.

## 4.5 Electrons

Electrons are reconstructed by matching tracks in the silicon tracker to the energy deposits in the ECAL [54]. The large amount of material in the CMS tracker causes electrons to emit bremsstrahlung photons along the trajectories before electrons reach the ECAL. As a result, an electron track often deviates more from a perfect helical trajectory compared to other charged particles. The energy deposits in the ECAL from the bremsstrahlung photons also spread along the $\phi$ direction due to the bending the electron tracks in the CMS magnetic field. Therefore, dedicated ECAL energy clustering algorithms as well as electron track reconstruction algorithms are necessary to precisely measure the energy and momentum of electrons.

In the barrel part of the ECAL (EB), a "hybrid" algorithm is used for electron energy clustering. The hybrid algorithm exploits the projective orientation of the crystals in the barrel. The crystal with the highest energy deposit in any considered region is selected as a seed. The hybrid algorithm starts by forming a $5\times1$ crystal array in $\eta\times\phi$ centered around the seed crystal. A cluster is formed by expanding the crystal array in both $\phi$ directions if the adjacent $5 \times 1$ crystal

array contains energy above a configurable threshold. The clusters are further collected into the final global cluster named supercluster (SC). In the ECAL endcap (EE), a "multi-$5 \times 5$" algorithm is used. The seed crystals in the endcaps are chosen as the ones with the largest energy deposits locally among their four directly adjoining crystals. Clusters of $5 \times 5$ crystals are formed around the seeds and are grouped into superclusters if they are within a configurable $\eta$-$\phi$ range and have minimum energy deposits. Energy deposits in the preshower that match to the SC extrapolations are also added to the total energy of the SC. The position of the SC is calculated as the mean position of its constituents weighted by the logarithm of the crystal energy.

Dedicated track reconstruction for electrons is initialized from seeds that are likely from electron trajectories. Such seeds are found by two complementary algorithms based on either the ECAL or the tracker. The ECAL-based seeding uses the SC energy and position to estimate the electron tracks in the inner layers of the silicon tracker. Seeds are selected from the hits in the corresponding layers if they are within $\phi$-$z$ (or $\phi$-$r$) windows of the SC extrapolation. The tracker-based seeding starts with tracks reconstructed as described in Section 4.2. These tracks are extrapolated to the ECAL and are matched to the PF clusters based on their positions and momenta. If a match is found, the seed of the matched track is used to seed electron track reconstruction in the next step. Otherwise, a second selection is tried among the unmatched tracks to recover potential electron tracks that have significant bremsstrahlung. Tracks with small number of hits or a large $\chi^2$ from the Kalman filter are refitted using a Gaussian sum filter (GSF) to account for the energy loss within the tracker volume. Compared to the ECAL-based seeding, the tracker-based seeding has higher efficiency for non-isolated electrons with low $p_{\mathrm{T}}$ or electrons in the barrel-endcap transition

region. The electron seeds selected from both methods are combined to initiate track reconstruction, which is based on the combinatorial Kalman filter method. The Bethe-Heitler function [55] is used to model the electron energy loss, and the match between predicted and reconstructed hits in a layer is less restrictive during the track building process. The final track parameters are extracted by a GSF fit. The GSF models the energy loss in each layer by a mixture of Gaussian distributions.

The electron candidates are built from combining the GSF tracks with the ECAL clusters. For ECAL-seeded electrons, the GSF tracks and the superclusters that seeded the track are combined with additional requirements that the GSF tracks are compatible with the SC positions in both $\eta$ and $\phi$. In case of tracker-seeded electrons, a loose selection is applied on a multivariate discriminant computed using the combined information on track reconstruction, shape and pattern of the PF cluster, and compatibility of the GSF track and the PF cluster.

Several electron identification methods are developed by CMS. In order to separate prompt and isolated electrons from misidentified jets, the $t\bar{t}H$ analysis presented in this thesis utilizes an identification strategy using multivariate analysis (MVA) based on boosted decision trees (BDT). The BDT-based algorithm improves the identification sensitivity by combining several discriminating observables including variables that characterize the shower shapes and SC substructure in the ECAL, fraction of energy deposited to HCAL and preshower, fraction of energy loss via bremsstrahlung, variables related to the fitted track, and compatibility of the track and ECAL SC. The performance of the electron BDT-based discriminants is shown in Fig. 4.4. The BDT is trained

with and without isolation variables. The performance of applying cuts on isolation variables on top of the BDT trained without the isolation variables are also shown in the plot. Electron candidates are required to have $p_T > 20$ GeV and $|\eta| < 2.5$. The performance is evaluated using simulated Drell-Yan samples under 2017 detector conditions. Signals (backgrounds) are reconstructed electrons that match (don't match) to genuine electrons at generator level within a cone of size $\Delta R = 0.1$. The electron MVA ID adopted in this analysis is the one

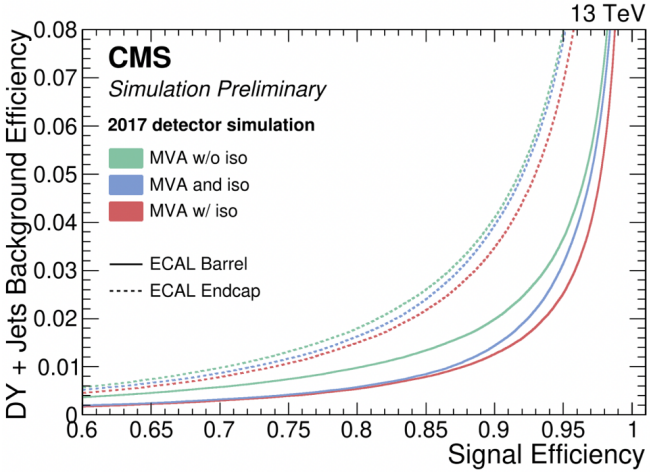


Figure 4.4: Receiver operating characteristic curves of the electron BDT-based identification [56].

without isolation variables. The isolation information is used later in the dedicated electron selection for this $t\bar{t}H$ analysis, which is discussed in Section 5.2.1. Extra identification criteria are applied to veto non-prompt electrons that are from a photon converting to an electron-positron pair. The photon conversion often happens in the tracker volume, and the secondary electrons tend to have missing hits in the innermost pixel layers. Therefore, an electron candidate is rejected if its track has more than one missing hits in the pixel detector, or if there exists a nearby electron track of opposite charge so that the pair can be fitted to a common vertex.

Like muons, an additional layer of dedicated electron selections, discussed in details in Section 5.2, is applied in the context of the $t\bar{t}H$ analysis presented in this thesis to further identify the signal electrons and facilitate the background estimation.

## 4.6 Jets

Jets are reconstructed by clustering the PF particle candidates using the anti-$k_t$ algorithm [51] with a distance parameter $R = 0.4$ implemented in FastJet [57]. In order to mitigate the effects of pileup, the PF candidates that do not originate from the primary vertex, as defined in Section 4.3, are excluded from the clustering.

Due to the non-linear detector response to particles, a set of corrections is applied to calibrate the reconstructed jet energy in order to better match the energy of the true particles or partons. The jet energy correction is based on a factorized approach, in which rescaling of the jet four-momentum is applied sequentially at different levels to deal with different effects. The first level or correction (L1 correction) aims to remove energy from pileup events, followed by other correction steps responsible for matching reconstructed jets in the simulation to the Monte Carlo truth particles and for correcting residual differences of jet response in collision data and simulation. The scale factors depend on pileup energy density, jet area, as well as jet $\eta$ and $p_{\mathrm{T}}$. In addition, the resolution of jet energy in data is found to be worse than that in the simulation. As a result, the jet energy and momentum in the simulation are smeared to better model the collision data. More details on the correction of jet energy scale and resolution

can be found in Ref. [58].

The PF jets used in this analysis are required to satisfy a set of identification requirement to suppress the fake ones due to calorimeter noise. The identification is based on the energy fractions as well as the multiplicity of different types of PF candidates clustered into the jet. Details of the identification criteria in different working points, namely loose and tight, are listed in Table 1 of Ref. [59]. The loose PF jet identification is used in this analysis using data collected in 2016. The analysis based on 2017 data uses the tight working point. The efficiencies of the PF jet identifications are above 99% in all $\eta$ range within the detector coverage [59].

## 4.7 b tagging

Due to the relatively long lifetime of b mesons ($\sim 1.5$ ps), jets that originate from b quark hadronization are usually displaced from the primary vertex by typically a few millimeters. So it is possible to reconstruct a secondary vertex using the displaced tracks from the b hadron decay, which can be used to tag jets from heavy quark decay. Two b tagging algorithms based on machine learning techniques are relevant to this analysis: CSVv2 and DeepCSV [60]. Both algorithms are based on the combined secondary vertex (CSV) taggers [61] used during LHC Run 1, exploiting the observables related to the properties as well as multiplicities of the reconstructed secondary vertices and the associated tracks. The CSVv2 tagger uses extra discriminating variables compared to the previous CSV tagger and improves the performance by training an artificial neutral network with one hidden layer. The DeepCSV tagger further improves the b tagging by

taking advantage of the rapid development in the field of deep learning and utilizes a deep neutral network for multi-label classification with more hidden layers and more nodes on each layer. Two working points loose and medium, corresponding to different thresholds on the tagger outputs, are used in this analysis. For both CSVv2 and DeepCSV taggers, the loose and medium working points have misidentification rates of around 10% and 1% respectively. The loose and medium working points of CSVv2 correspond to about 80% and 65% efficiency of tagging b jets with $p_{\mathrm{T}} > 20$ GeV, while for DeepCSV the corresponding efficiency is around 85% and 70%, respectively.

## 4.8 Hadronic taus

The tau ($\tau$) lepton has a mass of $1776.86 \pm 0.12$ MeV [62] and is the only lepton that is heavy enough to decay into hadrons. About one third of tau leptons decay into muons or electrons together with neutrinos. Muons and electrons are reconstructed as described in Section 4.4 and 4.5. The other two thirds of tau leptons decay into hadrons and a tau neutrino. While neutrinos escape the detector, the hadronic systems from tau decays, denoted as $\tau_{\mathrm{h}}$, are reconstructed by the "hadron-plus-strips" (HPS) algorithm [63]. Several major $\tau$ hadronic decay modes are reconstructed including: $\mathrm{h}^{\pm}$, $\mathrm{h}^{\pm}\pi^0$, $\mathrm{h}^{\pm}\pi^0\pi^0$, and $\mathrm{h}^{\pm}\mathrm{h}^{\mp}\mathrm{h}^{\pm}$, where h represents either a pion or kaon, and $\pi^0$ denotes a neutral pion.

The reconstruction of hadronic tau decays is seeded from PF jets. The HPS algorithm combines charged hadron constituents in the PF jet with "strips" from neutral pions. The neutral pion decays into two photons promptly. Each photon is likely to convert into a pair of electron and positron in the tracker material

before reaching the ECAL. As a result, the energy deposit in the ECAL due to neutral pions is likely to spread in a region narrow in $\eta$ and wide in $\phi$, referred to as a "strip". The strip is reconstructed by clustering the photon and electron PF candidates in the jet within a $\Delta\eta\times\Delta\phi$ window. Both $\Delta\eta$ and $\Delta\phi$ are parameterized by the $p_{\mathrm{T}}$ of the clustered PF candidates, varying from 0.05 to 0.15 and from 0.05 to 0.3 respectively.[1] The momentum of the strip is computed as the vectorial sum of all its constituent momenta, and its $\eta$-$\phi$ coordinates are determined by a $p_{\mathrm{T}}$-weighted average of its constituents.

The charged hadron candidates used in the HPS algorithm are required to be compatible with originating from the primary vertex. They are combined with the strips in the jet under different $\tau$ decay mode hypotheses. The mass of the reconstructed $\tau_{\mathrm{h}}$ candidate is required to be compatible with the mass of $\rho(770)$ meson if it is reconstructed in the $\mathrm{h}^{\pm}\pi^0$ mode, or with the mass of $\mathrm{a}_1(1260)$ meson if it is reconstructed in the $\mathrm{h}^{\pm}\pi^0\pi^0$ or $\mathrm{h}^{\pm}\mathrm{h}^{\mp}\mathrm{h}^{\pm}$ decay mode. The total charge of the $\tau_{\mathrm{h}}$ candidate has to be $\pm1$. Furthermore, a signal cone is established around the $\tau_{\mathrm{h}}$ candidate with $R_{\mathrm{sig}} = (3.0\ \mathrm{GeV})/p_{\mathrm{T}}$, limited to the range between 0.05 and 0.10. The $\tau_{\mathrm{h}}$ candidate is rejected if any of its constituent charged hadrons or strips are outside the signal cone.

Energies of reconstructed $\tau_{\mathrm{h}}$'s in simulation are calibrated to better match those measured in recorded data. Scale factors for such corrections applied to $\tau_{\mathrm{h}}$'s in simulation are determined in $\mathrm{Z}/\gamma^* \rightarrow \tau\tau$ events, in which one $\tau$ decays into a lepton and the other $\tau$ decays hadronically. The scale factors are extracted separately in each $\tau_{\mathrm{h}}$ decay mode by fitting distributions of the reconstructed $\tau_{\mathrm{h}}$ mass as well as the mass of the lepton and $\tau_{\mathrm{h}}$ pair. These two variables are

[1] A fixed window size of $\Delta\eta = 0.05$ and $\Delta\phi = 0.20$ is used in the previous version of the HPS algorithm [64] for the LHC Run 1.

useful for this measurement since they are sensitive to the $\tau_h$ energy scale. More information on the $\tau$ energy scale measurement can be found in Ref. [63].

Isolation requirements are applied to the $\tau_h$ candidates to better distinguish them from light quark and gluon jets. The isolation is calculated as a sum of scalar $p_T$ of the charged PF candidates and PF photons within an isolation cone around the $\tau_h$ candidate, excluding those used in the $\tau_h$ reconstruction. The size of the isolation cone used in this analysis has a radius $R_{iso}$ of 0.3, which is smaller than the cone size of 0.5 adopted by most of other CMS analyses. This smaller isolation cone is optimized for the final states with higher hadronic activity as in $t\bar{t}H$ events. The isolation information, the decay mode, the multiplicity of photon and electron candidates, and variables sensitive to the $\tau$ lifetime are used as input variables to a classifier based on boosted decision trees to achieve a better discrimination between $\tau_h$ and jets. The BDT is trained on Monte Carlo samples of hadronic tau decays in $t\bar{t}H$ events as signal and jets in $t\bar{t}$ + jets as background.

## 4.9 Missing transverse momentum

Particles that do not interact with the detector, for example neutrinos, would leave a deficiency of the total momentum in the transverse plane. The missing transverse momentum, denoted as $\vec{p}_T^{\,miss}$, is computed in each event as the negative vectorial sum of the transverse momenta of all PF candidates [65]. The energy correction of the PF jets mentioned above are propagated to the calculation of $\vec{p}_T^{\,miss}$, which improves the resolution and response of the $\vec{p}_T^{\,miss}$ measurement. The magnitude of $\vec{p}_T^{\,miss}$ is denoted as $p_T^{miss}$ or $E_T^{miss}$.

CHAPTER 5

# MEASUREMENT OF $t\bar{t}H$ PRODUCTION CROSS SECTION IN FINAL STATES WITH LEPTONS AND HADRONIC TAUS

As discussed previously in Section 2.4.3, measurements of the $t\bar{t}H$ cross section provide direct constraints on the top-Higgs coupling strength. Both the top quark and the Higgs boson have various decay modes, leading to a rich set of final states with distinctive features. This chapter describes the $t\bar{t}H$ analysis in final states with electrons, muons, and hadronic taus. Such final states target the decay modes in which at least one of the top quarks decay leptonically, and the Higgs boson decays into a pair of gauge bosons (W or Z) or a pair of $\tau$ leptons. Data recorded by the CMS experiment in both 2016 and 2017 during LHC Run 2 with proton-proton collisions at a center-of-mass energy of 13 TeV are used. The result based on the 2016 data set contributed to the combination of all CMS $t\bar{t}H$ efforts that lead to the first observation of the $t\bar{t}H$ production at the LHC [17]. The analysis using the 2017 data follows the similar strategy as the 2016 version with a few improvements. Discussions in the rest of this chapter are based on the 2017 version of the analysis with the intention to provide a self-consistent narrative as well as most up-to-date information, while still convey the general strategy adopted in this $t\bar{t}H$ analysis channel.

This chapter is organized as follows: data sets as well as Monte Carlo samples used in the analysis are described in Section 5.1. Section 5.2 discusses in details about the dedicated physics object selections used in this analysis. Event selection criteria are presented in Section 5.3, followed by discussions on signal and background estimations in Section 5.4. Discriminating observables as well as the statistical method used for extracting the $t\bar{t}H$ signal are described in Sec-

tion 5.5. Section 5.6 summaries relevant systematic uncertainties. Finally, the analysis result is presented in Section 5.7.

## 5.1 Data sets and Monte Carlo samples

The analysis presented in this chapter are based on data collected in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2017, corresponding to an integrated luminosity of 41.5 $\text{fb}^{-1}$. The events are recorded using single, double, and triple lepton triggers, as well as triggers requiring a lepton (e or $\mu$) and a $\tau_\text{h}$. Recorded events are stored in several data sets based on the trigger paths and data taking periods. Each event is identified uniquely by its run number, luminosity section (LS), and the event index in its LS. A dedicated team in CMS validates the quality of the recorded data and provides a list of luminosity sections that are not marked as data of good quality due to e.g. detector operation conditions. These data are masked and are not used in this analysis. A detailed list of data sets used in the analysis, together with their corresponding run number range as well as the integrated luminosities, is shown in Table 5.1. Duplicate data events stored in multiple primary data sets are removed based on their event identifiers.

Samples generated by Monte Carlo (MC) simulations are used to model both the $\text{t}\bar{\text{t}}\text{H}$ signal and relevant background processes. The Higgs boson mass is assumed to be 125 GeV, and the top quark mass is set to 172.5 GeV for all simulated processes. The signal MC samples are produced using next-to-leading-order (NLO) matrix elements implemented in the MADGRAPH5_aMC@NLO program [66–69]. One of the main backgrounds arises from the production of a top quark pair associated with a W or Z boson ($\text{t}\bar{\text{t}}\text{W}$, $\text{t}\bar{\text{t}}\text{Z}$). Both processes are generated at NLO accuracy by MADGRAPH5_aMC@NLO. The process of a top quark pair associated with two W bosons ($\text{t}\bar{\text{t}}\text{WW}$) also has a sizable contribution to the total background and is simulated at leading-order (LO) accuracy. Other

Table 5.1: List of data sets in 2017 used in this analysis.

| Data set name | Run-range | Luminosity ($fb^{-1}$) |
|---|---|---|
| /SingleElectron/Run2017B-17Nov2017-v1 | 297047–299329 | 4.79 |
| /SingleElectron/Run2017C-17Nov2017-v1 | 299368–302029 | 9.63 |
| /SingleElectron/Run2017D-17Nov2017-v1 | 302030–302663 | 4.25 |
| /SingleElectron/Run2017E-17Nov2017-v1 | 303818–304797 | 9.31 |
| /SingleElectron/Run2017F-17Nov2017-v1 | 305040–306460 | 13.54 |
| /SingleMuon/Run2017B-17Nov2017-v1 | 297047–299329 | 4.79 |
| /SingleMuon/Run2017C-17Nov2017-v1 | 299368–302029 | 9.63 |
| /SingleMuon/Run2017D-17Nov2017-v1 | 302031–302663 | 4.25 |
| /SingleMuon/Run2017E-17Nov2017-v1 | 303824–304797 | 9.31 |
| /SingleMuon/Run2017F-17Nov2017-v1 | 305040–306462 | 13.54 |
| /DoubleEG/Run2017B-17Nov2017-v1 | 297047–299329 | 4.79 |
| /DoubleEG/Run2017C-17Nov2017-v1 | 299368–302029 | 9.63 |
| /DoubleEG/Run2017D-17Nov2017-v1 | 302030–302663 | 4.25 |
| /DoubleEG/Run2017E-17Nov2017-v1 | 303818–304797 | 9.31 |
| /DoubleEG/Run2017F-17Nov2017-v1 | 305040–306460 | 13.54 |
| /DoubleMuon/Run2017B-17Nov2017-v1 | 297047–299329 | 4.79 |
| /DoubleMuon/Run2017C-17Nov2017-v1 | 299368–302029 | 9.63 |
| /DoubleMuon/Run2017D-17Nov2017-v1 | 302031–302663 | 4.25 |
| /DoubleMuon/Run2017E-17Nov2017-v1 | 303824–304797 | 9.31 |
| /DoubleMuon/Run2017F-17Nov2017-v1 | 305040–306462 | 13.54 |
| /MuonEG/Run2017B-17Nov2017-v1 | 297047–299329 | 4.79 |
| /MuonEG/Run2017C-17Nov2017-v1 | 299368–302029 | 9.63 |
| /MuonEG/Run2017D-17Nov2017-v1 | 302031–302663 | 4.25 |
| /MuonEG/Run2017E-17Nov2017-v1 | 303824–304797 | 9.31 |
| /MuonEG/Run2017F-17Nov2017-v1 | 305040–306460 | 13.54 |

simulated backgrounds include diboson processes (WZ, ZZ, WW), a pair of top quarks associated with a real or virtual photon ($t\bar{t}\gamma$, $t\bar{t}\gamma^*$), single top associated with a photon, and other rare SM processes such as tribosons (WWW, WWZ, WZZ, ZZZ), single top quark associated with a Z boson (tZq), same-sign WW, WW double scattering, and $t\bar{t}t\bar{t}$. These background processes are generated by either MADGRAPH5_aMC@NLO or POWHEG [70–76] at NLO, except that the WW ($2\ell 2\nu$) sample is produced with its cross section calculated at next-to-next-to-leading-order (NNLO) accuracy, and the same-sign WW and WW double scattering processes are generated at LO accuracy. The Higgs boson production

associated with a single top quark (tH) is treated as a background in this analysis. The tH sample is generated at LO by MADGRAPH5 with its cross section set to the SM prediction, assuming a destructive interference between the process in which the Higgs boson couples to a top quark and the process in which H couples to a W boson.

Samples of the top quark pair ($t\bar{t}$) production, the single top quark production, W+jets, and the Drell-Yan process are also generated for validating the date-driven background estimations and for other auxiliary measurements. The production of top quark pairs are simulated by POWHEG at NLO accuracy, and the single top sample is generated at NLO by MADGRAPH5_aMC@NLO. The W+jets and Drell-Yan ($Z/\gamma^* \to \ell\ell$) samples are normalized to their NNLO cross sections and are produced by MADGRAPH5. Both the signal and background samples are generated using the NNPDF3.0 [77] set of parton distribution functions (PDFs).

The parton shower and hadronization processes in all samples are modeled by PYTHIA 8.212 [78] with the CP1-5 tune [79]. PYTHIA is also used to model the $\tau$ lepton decay, including the polarization effect.

Minimum bias events generated by PYTHIA, assuming a proton-proton inelastic cross section of 69.2 mb, are overlaid on all MC samples to mimic the pileup condition in the recorded data.

The response of MC particles going through the full CMS detector is simulated using GEANT4 [80]. All simulated events are processed using the same particle reconstruction and identification software as used for processing collision data.

## 5.2 Object selection

The reconstruction and basic identification of physics objects used in this analysis are describe in Chapter 4. Additional object selection criteria optimized for this analysis are detailed in this section.

### 5.2.1 Muons and electrons

The basic identifications described in Section 4.4 and 4.5 for muons and electrons are effective in separating genuine muons or electrons from jets. Dedicated lepton identification criteria are developed in the context of the $t\bar{t}H$ analysis to further distinguish between the "prompt" leptons from W, Z, or $\tau$ decays and the "non-prompt" ones produced in the decays of heavy (charm or bottom) quarks.

One distinctive feature between the prompt and non-prompt leptons is their level of isolation. A prompt lepton is expected to be isolated, while a non-prompt lepton is usually accompanied by other hadronic activities in its vicinity. To quantify the level of isolation of a given lepton, an observable $I^{\ell}$ is defined as

$$I^{\ell} = \sum_{\text{charged}} p_{\text{T}} + \max\left[0, \sum_{\text{neutrals}} p_{\text{T}} - \rho\,\mathcal{A}\left(\frac{R}{0.3}\right)^2\right]. \tag{5.1}$$

The first term, denoted as $I^{\ell}_{\text{charged}}$, is the scalar $p_{\text{T}}$ sum of all charged particles within a cone of size $R$ centered around the lepton. All charged particles entering the sum are required to originate from the same collision vertex as the lepton in order to mitigate the pileup effect. The cone size $R$, defined as $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, varies between 0.05 and 0.20 and is proportional to the in-

verse of the lepton $p_T$:

$$R = \begin{cases} 0.05 & \text{if } p_T > 200 \text{ GeV} \\ 10 \text{ GeV}/p_T & \text{if } 50 < p_T < 200 \text{ GeV} \\ 0.20 & \text{if } p_T < 50 \text{ GeV} \end{cases} \quad (5.2)$$

The smaller cone size for larger lepton $p_T$ not only reduces the pileup effect but also gives higher lepton identification efficiency in the $t\bar{t}H$ signal events with high hadronic activities. The second term in Eq. 5.1, denoted as $I^{\ell}_{\text{neutral}}$, represents the neutral component of the isolation and is computed as the scalar $p_T$ sum of the neutral hadrons and photons within the cone. This sum is corrected by subtracting extra contributions due to pileup. The neutral pileup correction is estimated by $\rho \mathcal{A} \left(\frac{R}{0.3}\right)^2$, in which $\rho$ is the energy density of neutral particles within the tracker acceptance, and $\mathcal{A}$ is the effective area. The neutral energy density $\rho$ is computed as described in Refs. [81,82]. The effective area $\mathcal{A}$ is given in Table 5.2 separately for muons and electrons and in bins of $\eta$. The values in the table are determined based on the correlation between $I^{\ell}$ and $\rho$ in simulations.

Table 5.2: Effective areas $\mathcal{A}$ for muons (left) and electrons (right).

| Muons | |
|---|---|
| Pseudorapidity range | $\mathcal{A}$ |
| $0.0 < \|\eta\| < 0.8$ | 0.0566 |
| $0.8 < \|\eta\| < 1.3$ | 0.0562 |
| $1.3 < \|\eta\| < 2.0$ | 0.0363 |
| $2.0 < \|\eta\| < 2.2$ | 0.0119 |
| $2.2 < \|\eta\| < 2.5$ | 0.0064 |

| Electrons | |
|---|---|
| Pseudorapidity range | $\mathcal{A}$ |
| $0.0 < \|\eta\| < 1.0$ | 0.1566 |
| $1.0 < \|\eta\| < 1.479$ | 0.1626 |
| $1.479 < \|\eta\| < 2.0$ | 0.1073 |
| $2.0 < \|\eta\| < 2.2$ | 0.0854 |
| $2.2 < \|\eta\| < 2.3$ | 0.1051 |
| $2.3 < \|\eta\| < 2.4$ | 0.1204 |
| $2.4 < \|\eta\| < 2.5$ | 0.1524 |

Besides the isolation, discriminating information can be obtained from the nearest jet to the lepton under study. For each reconstructed lepton, several ob-

servables related to its nearest matched jet are calculated, including $p_{\mathrm{T}}^{\ell}/p_{\mathrm{T}}^{\mathrm{jet}}$, the $p_{\mathrm{T}}$ ratio of the lepton and the jet, where $p_{\mathrm{T}}^{\ell}$ and $p_{\mathrm{T}}^{\mathrm{jet}}$ denote the transverse momentum of the lepton and its nearest jet, respectively; $p_{\mathrm{T}}^{\mathrm{rel}}$, the component of lepton momentum that is perpendicular to the jet momentum ("relative" $p_{\mathrm{T}}$); $N_{\mathrm{charged}}^{\mathrm{jet}}$, the number of charged constituents of the jet; jetBTag, the discriminant value of the b tagging algorithm (DeepCSV) of the jet. The momentum and energy of the jet in these calculation are corrected based on the so-called "lepton-aware" jet energy correction in order to avoid over-correcting the lepton energy. After applying the first level of energy correction to the raw jet (c.f. Section 4.6), the lepton is subtracted from the jet before applying the rest of energy corrections. The lepton momentum is added back to the total jet momentum in the end to get the corrected jet four-momentum. The matched jet is required to be within a distance $\Delta R < 0.4$ to the lepton and has a $p_{\mathrm{T}}$ greater than 15 GeV. In case there is no such jet found, $p_{\mathrm{T}}^{\ell}/p_{\mathrm{T}}^{\mathrm{jet}}$ is set to $p_{\mathrm{T}}^{\ell}/(p_{\mathrm{T}}^{\ell} + I^{\ell})$, and other variables are set to zero.

The discriminating power of individual observables are enhanced by a BDT classifier trained separately for muons and electrons, referred to as prompt-$\mu$ or prompt-e BDT. The input variables of the BDT includes:

- $p_{\mathrm{T}}$ and $\eta$ of the lepton
- Isolation variables: $I_{\mathrm{charged}}^{\ell}$ and $I_{\mathrm{neutral}}^{\ell}$
- Variables related to the nearest jet: $p_{\mathrm{T}}^{\ell}/p_{\mathrm{T}}^{\mathrm{jet}}$, $p_{\mathrm{T}}^{\mathrm{rel}}$, $N_{\mathrm{charged}}^{\mathrm{jet}}$, and jetBTag
- Impact parameters of the lepton: $d_{xy}$ and $d_z$ (the transverse and longitudinal impact parameters of the lepton track), $d/\sigma_d$ (the significance of the 3D impact parameter)

- Basic object ID to discriminate against jets: muon segment compatibility, as described in Section 4.4, or electron non-triggering ID based on multivariate analysis, referred to as electron MVA ID hereafter, as discussed in Section 4.5.

The prompt-$\mu$ and prompt-e BDTs are trained on samples of prompt leptons in simulated $t\bar{t}H$ events as signals and non-prompt leptons in simulated $t\bar{t}$ + jets events as backgrounds.

Three lepton selection working points (WPs) are defined for muons and electrons: loose, fakeable, and tight. The tight leptons are the nominal collection used to select events in signal regions. The fakeable leptons have relaxed selection requirements compared to the tight leptons and are used for background estimations as discussed in Section 5.4.1. The loose leptons are selected with basic kinematic, isolation, and identification requirements and are used to define lepton collections for training. The loose leptons are also used in the calculations for mass veto in event selections described in Section 5.3.7. The selection criteria for the three working points are detailed in Table 5.3 for muons and Table 5.4 for electrons. The loose working point of the electron MVA ID is $p_{\mathrm{T}}$ and $\eta$ dependent and is defined in Table 5.5.

For electrons, an additional set of selections is applied to mimic the electron identification requirements applied at the trigger level. The selection criteria involve the width of the electron cluster in $\eta$ ($\sigma_{i\eta i\eta}$), the ratio of the HCAL energy deposit to the ECAL energy deposit associated with the electron (H/E), and the difference between the inverse of the electron cluster energy and the inverse of its track momentum (1/E - 1/p).

Table 5.3: Loose, fakeable, and tight selection criteria for muons. A hyphen (−) indicates the selection is not applied. The relaxed selection criteria for the fakeable selection in case of prompt-$\mu$ BDT > 0.90 are placed in parentheses.

| Observable | Loose | Fakeable | Tight |
|---|---|---|---|
| $p_{\mathrm{T}}$ † | > 5 GeV | > 10 GeV | > 10 GeV |
| $\lvert\eta\rvert$ | < 2.4 | < 2.4 | < 2.4 |
| $\lvert d_{xy}\rvert$ | < 0.05 cm | < 0.05 cm | < 0.05 cm |
| $\lvert d_z\rvert$ | < 0.1 cm | < 0.1 cm | < 0.1 cm |
| $d/\sigma_d$ | < 8 | < 8 | < 8 |
| $I_\mu$ | $< 0.4 \times p_{\mathrm{T}}$ | $< 0.4 \times p_{\mathrm{T}}$ | $< 0.4 \times p_{\mathrm{T}}$ |
| Loose PF muon | ✓ | ✓ | ✓ |
| Medium PF muon | – | – | ✓ |
| Segment compatibility | – | > 0.3 (–) | – |
| $p_{\mathrm{T}}^{\mu}/p_{\mathrm{T}}^{\mathrm{jet}}$ | – | > 0.6 (–) | – |
| DeepCSV of the nearest jet | – | < 0.07 (< 0.4941) | < 0.4941 |
| Prompt-$\mu$ BDT | – | – | > 0.90 |

† cone-$p_{\mathrm{T}}$ for fakeable and tight selections

As indicated in Table 5.3 and 5.4, prompt BDT output > 0.90 is required for tight selections, while no such requirement is applied for loose or fakeable leptons. For fakeable selections, however, if the lepton's prompt BDT output is greater than 0.90, the selection criteria on $p_{\mathrm{T}}^{\ell}/p_{\mathrm{T}}^{\mathrm{jet}}$ and DeepCSV of the nearest jet, as well as the segment compatibility in case of muons or the electron MVA ID in case of electrons, are relaxed to the same conditions as the tight selections. This ensures the tight lepton collection is strictly a subset of the fakeable collection. The relaxed criteria for fakeable selections in case of prompt BDT > 0.90 are placed in parentheses in Table 5.3 and 5.4.

For fakeable and tight leptons, a variable "cone-$p_{\mathrm{T}}$" is defined in order to reduce biases in background estimations. The cone-$p_{\mathrm{T}}$ is set to the lepton $p_{\mathrm{T}}$ if it passes the tight selection. In case a lepton passes the fakeable selection criteria but fails tight, the cone-$p_{\mathrm{T}}$ of the lepton is set to 0.9 times the $p_{\mathrm{T}}^{\mathrm{jet}}$ of its matched jet if the distance $\Delta R$ between the lepton and the jet is less than 0.4, otherwise

Table 5.4: Loose, fakeable, and tight selection criteria for electrons. A hyphen (−) indicates the selection is not applied. The relaxed selection criteria for the fakeable selection in case of prompt-$\mu$ BDT $>$ 0.90 are placed in parentheses. The requirements on $\sigma_{i\eta i\eta}$ are different in the barrel and endcap regions. The numbers in the curly brackets are separated by | and refer to $|\eta| < 1.479$ and $|\eta| \geq 1.479$, respectively.

| Observable | Loose | Fakeable | Tight |
|---|---|---|---|
| $p_{\mathrm{T}}$ [†] | $> 7$ GeV | $> 10$ GeV | $> 10$ GeV |
| $\lvert\eta\rvert$ | $< 2.5$ | $< 2.5$ | $< 2.5$ |
| $\lvert d_{xy}\rvert$ | $< 0.05$ cm | $< 0.05$ cm | $< 0.05$ cm |
| $\lvert d_z\rvert$ | $< 0.1$ cm | $< 0.1$ cm | $< 0.1$ cm |
| $d/\sigma_d$ | $< 8$ | $< 8$ | $< 8$ |
| $I_{\mathrm{e}}$ | $< 0.4 \times p_{\mathrm{T}}$ | $< 0.4 \times p_{\mathrm{T}}$ | $< 0.4 \times p_{\mathrm{T}}$ |
| Electron MVA ID | Loose WP | $> 0.50$ (Loose WP) | Loose WP |
| $\sigma_{i\eta i\eta}$ | – | $< \{0.011 \mid 0.030\}$ | $< \{0.011 \mid 0.030\}$ |
| H/E | – | $< 0.10$ | $< 0.10$ |
| 1/E - 1/p | – | $> -0.04$ | $> -0.04$ |
| Photon conversion veto | – | ✓ | ✓ |
| Missing hits | $\leq 1$ | $= 0$ | $= 0$ |
| $p_{\mathrm{T}}^{\mathrm{e}}/p_{\mathrm{T}}^{\mathrm{jet}}$ | – | $> 0.6$ (–) | – |
| DeepCSV of the nearest jet | – | $< 0.07$ ($< 0.4941$) | $< 0.4941$ |
| Prompt-e MVA | – | – | $> 0.90$ |

[†] cone-$p_{\mathrm{T}}$ for fakeable and tight selections

Table 5.5: Definition of the electron MVA ID loose working point.

| Electron MVA Loose ID | $\lvert\eta\rvert < 0.80$ | $0.80 \leq \lvert\eta\rvert < 1.479$ | $\lvert\eta\rvert \geq 1.479$ |
|---|---|---|---|
| $p_{\mathrm{T}} < 10$ GeV | $> -0.13$ | $> -0.32$ | $> -0.08$ |
| $p_{\mathrm{T}} \geq 10$ GeV | $> -0.86$ | $> -0.81$ | $> -0.72$ |

the cone-$p_{\mathrm{T}}$ is set to 0.9 times $p_{\mathrm{T}}^{\ell} + I^{\ell}$, where $p_{\mathrm{T}}^{\ell}$ is the lepton's $p_{\mathrm{T}}$ and $I^{\ell}$ is the lepton's isolation as defined in Eq. 5.1.

It is worth mentioning that the prompt lepton identification described in this section targets primarily the electrons and muons originating from W or Z bosons. Given the sizable lifetime ($c\tau = 87$ $\mu$m) of $\tau$ leptons, the electrons and muons from $\tau$ lepton decays tend to be less "signal-like" especially in distributions of $d_{xy}$, $d_z$, and $d/\sigma_d$. As a result, the prompt lepton identification is slightly

less optimal for leptons from $\tau$ decays. More sophisticated algorithms for multi-classification could be useful to improve the lepton identification efficiency in the future version of the analysis.

### 5.2.2 Hadronic taus

As described in Section 4.8, hadronic taus are reconstructed by the HPS algorithm and are identified using a smaller isolation cone with a radius $R_{\text{iso}}$ of 0.3. A multivariate discriminant based on the BDT, referred to as $\tau_{\text{h}}$ MVA ID, is used to distinguish between hadronic taus and jets. Various working points are defined based on different thresholds of the BDT output. The working points used in this analysis include very-loose, loose, and medium, corresponding to a $\tau_{\text{h}}$ identification efficiency (jet misidentification rate) of approximately 60% (2%), 55% (0.7%), and 50% (0.4%), respectively.

Two levels of $\tau_{\text{h}}$ selections are used in this analysis. The selection criteria are listed in Table 5.6. Similar to muons and electrons, the tight selection for $\tau_{\text{h}}$ is used for selecting events in signal regions, and the fakeable $\tau_{\text{h}}$ selection is used for background estimations discussed in details in Section 5.4.1. The "old" decay mode in Table 5.6 refers one of the main $\tau_{\text{h}}$ decay modes: $\text{h}^{\pm}$, $\text{h}^{\pm}\pi^0$, $\text{h}^{\pm}\pi^0\pi^0$, or $\text{h}^{\pm}\text{h}^{\mp}\text{h}^{\pm}$, where h denotes either a pion or kaon and $\pi^0$ stands for a neutral pion. For the tight selection, the loose WP of $\tau_{\text{h}}$ MVA ID is used in the event categories with exactly one $\tau_{\text{h}}$. In case of the event categories with at least two $\tau_{\text{h}}$'s, the $\tau_{\text{h}}$ MVA ID requirement is tightened to the medium WP. Event categorizations based on multiplicities of leptons and hadronic taus are described in Section 5.3.

Table 5.6: Fakeable, and tight selection criteria for hadronic taus.

| Observable | Fakeable | Tight |
|---|---|---|
| $p_T$ | $> 20$ GeV | $> 20$ GeV |
| $\|\eta\|$ | $< 2.3$ | $< 2.3$ |
| $\|d_z\|$ | $< 0.2$ cm | $< 0.2$ cm |
| Decay mode | old | old |
| $\tau_h$ MVA ID ($R = 0.3$) | Very-loose WP | Loose or Medium WP |

No dedicated discriminators are used to separate hadronic taus from misidentified electrons or muons. Instead, any $\tau_h$ candidates that are within $\Delta R = 0.3$ of the loose electrons or muons are removed from both the fakeable and tight $\tau_h$ collections.

### 5.2.3 Jets and missing transverse momentum

Jets are reconstructed as described in Section 4.6. The jets used in this analysis are required to satisfy $p_T > 25$ GeV and $|\eta| < 2.4$. Any jet within a distance of $\Delta R = 0.4$ to any fakeable electrons, muons, or hadronic taus are not selected. The DeepCSV algorithm described in Section 4.7 is adopted for tagging jets originating from b quarks in this version of the analysis.

In $t\bar{t}H$ signal events, neutrinos can be produced from decays of vector bosons or $\tau$ leptons, escaping the CMS apparatus without being detected. Their transverse momentum can be estimated by the missing transverse momentum $\vec{p}_T^{\,miss}$ as mentioned in Section 4.9. In addition, an observable $\vec{H}_T^{miss}$ is defined as the negative vectorial sum of the transverse momentum of selected jets as well as the electrons, muons, and $\tau_h$'s that pass the fakeable selections. The magnitude of $\vec{H}_T^{miss}$ ($|\vec{H}_T^{miss}|$) has worse resolution than $|\vec{p}_T^{\,miss}|$ but is more robust against

pileup, since the pileup hadrons are usually soft (i.e. of low $p_T$) and do not enter the calculation of $\vec{H}_T^{miss}$. In an event with a real missing transverse momentum caused by neutrinos, the $|\vec{H}_T^{miss}|$ and $|\vec{p}_T{}^{miss}|$ tend to be more correlated than in an event with the imbalance of total transverse momentum due to instrumental effects. Therefore, a discriminator $L_D$ is defined in Eq. 5.3 as a linear combination of $|\vec{H}_T^{miss}|$ and $|\vec{p}_T{}^{miss}|$ and is used in event selections to better identify events in which neutrinos are produced.

$$L_D = 0.6 \times |\vec{p}_T{}^{miss}| + 0.4 \times |\vec{H}_T^{miss}| \tag{5.3}$$

The two coefficients in the definition are optimized using simulated $t\bar{t}H$ and Drell-Yan events to provide the best discriminating power.

## 5.3 Event selection

The $t\bar{t}H$ process has a rich set of final states due to various decay channels of top quarks and the Higgs boson. Events are grouped into mutually exclusive categories based on numbers of selected leptons (e or $\mu$) and hadronic taus ($\tau_h$) that pass tight selections as detailed in Table 5.3, 5.4, and 5.6. Seven event categories are considered in this analysis:

- $1\ell + 2\tau_h$: one lepton and two $\tau_h$'s, targeting semi-leptonic $t\bar{t}$ and $H \to \tau\tau$, in which both $\tau$ leptons decay hadronically.
- $2\ell ss$: two leptons with same sign of charge and no $\tau_h$, targeting semi-leptonic $t\bar{t}$ and $H \to WW$, in which one of the W boson decays leptonically while the other decays hadronically.
- $2\ell ss + 1\tau_h$: two same-sign leptons and one $\tau_h$, targeting semi-leptonic $t\bar{t}$ and $H \to \tau\tau$, in which one $\tau$ decays leptonically.
- $2\ell + 2\tau_h$: two leptons and two $\tau_h$'s, targeting leptonic $t\bar{t}$ and $H \to \tau\tau$, in which both $\tau$ leptons decay hadronically.
- $3\ell$: three leptons and no $\tau_h$, targeting leptonic (semi-leptonic) $t\bar{t}$ and $H \to WW$, in which one W decays (both W's decay) leptonically.
- $3\ell + 1\tau_h$: three leptons and one $\tau_h$, targeting leptonic $t\bar{t}$ and $H \to \tau\tau$, in which one $\tau$ decays leptonically and the other decays hadronically.
- $4\ell$: four leptons, targeting leptonic $t\bar{t}$ and $H \to WW$, in which both W bosons decay leptonically, or $H \to ZZ$, in which one Z boson decays into a pair of leptons.

Event selection criteria are optimized in each event category to select $t\bar{t}H$ signal events and to suppress the relevant backgrounds. The set of selection criteria are described below. Event selection requirements applied in each category, as well as the targeted decay channels, are summarized in Table 5.8 and 5.9.

### 5.3.1 Triggers

A combination of triggers requiring the presence of leptons and $\tau_h$'s is used to record data for this analysis. The same trigger requirements are applied to the simulated events. Table 5.7 shows the types of triggers as well as the specific HLT paths used in this analysis. Events in $2\ell$ss, $2\ell\text{ss} + 1\tau_h$, and $2\ell + 2\tau_h$ categories

Table 5.7: High-level triggers used to record events selected in the $1\ell + 2\tau_h$, $2\ell$ss, $2\ell\text{ss} + 1\tau_h$, $2\ell + 2\tau_h$, $3\ell$, $3\ell + 1\tau_h$, and $4\ell$ channels.

| | |
|---|---|
| Single-lepton triggers | HLT_Ele32_WPTight_Gsf |
| | HLT_Ele35_WPTight_Gsf |
| | HLT_IsoMu24 |
| | HLT_IsoMu27 |
| Double-lepton triggers | HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL |
| | HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ |
| | HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL |
| | HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ |
| | HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ |
| | HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ |
| | HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8 |
| Triple-lepton triggers | HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL |
| | HLT_Mu8_DiEle12_CaloIdL_TrackIdL |
| | HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ |
| | HLT_TripleMu_12_10_5 |
| Lepton+$\tau_h$ cross triggers | HLT_Ele24_eta2p1_WPTight_Gsf_LooseChargedIsoPFTau30_eta2p1_CrossL1 |
| | HLT_IsoMu20_eta2p1_LooseChargedIsoPFTau27_eta2p1_CrossL1 |

are required to pass at least one of the HLT paths in either single-lepton triggers or double-lepton triggers. In case of $3\ell$, $3\ell + 1\tau_h$, and $4\ell$ categories, the collection of HLT paths are extended to triple-lepton triggers to increase the overall selection efficiency. In the $1\ell + 2\tau_h$ category, triggers that select events contain-

ing an electron or muon and a hadronic $\tau$ (lepton+$\tau_h$ cross triggers) are used in addition to single-lepton triggers.

### 5.3.2 Lepton and $\tau_h$ multiplicity

The requirement on numbers on leptons and $\tau_h$'s defines each event category. For signal region selections, exactly two tight leptons (e or $\mu$) are required in $2\ell$ss and $2\ell$ss+$1\tau_h$ categories. Similarly, $3\ell$ and $3\ell$+$1\tau_h$ categories require exactly three tight leptons, and exactly one tight lepton is required in the $1\ell + 2\tau_h$ category. For event category $2\ell + 2\tau_h$, two tight leptons are needed, but an event is still accepted in this category in case there are more tight leptons. Any events with four or more tight leptons are grouped into the $4\ell$ category.

As mentioned in Section 5.2.2, the tau MVA ID working point for tight $\tau_h$ selection varies in different event categories. A tighter tau MVA ID WP is used in event categories with more than one $\tau_h$ to further suppress backgrounds due to misidentification of jets as $\tau_h$'s. For $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$, at least two tight $\tau_h$'s passing the medium WP are required. $2\ell$ss and $3\ell$ categories veto events with any $\tau_h$ candidates passing Loose WP, while $2\ell$ss + $1\tau_h$ and $3\ell$ + $1\tau_h$ categories require at least one $\tau_h$ candidate passing loose WP. In addition, events selected in the $2\ell$ss+$1\tau_h$ category are also required to have no more than one $\tau_h$'s candidates passing medium WP in order to avoid double counting events that would be selected in the $2\ell + 2\tau_h$ category.

Table 5.8: Event selections applied in the $2\ell$ss, $2\ell$ss + $1\tau_h$, $1\ell$ + $2\tau_h$ and $2\ell$ + $2\tau_h$ categories.

| Selection | $2\ell$ss | $2\ell$ss + $1\tau_h$ |
|---|---|---|
| Targeted $t\bar{t}H$ decays | $t \to b\ell\nu$, $t \to bqq$, $H \to WW \to \ell\nu qq$ | $t \to b\ell\nu$, $t \to bqq$, $H \to \tau\tau \to \ell\tau_h + \nu$'s |
| Trigger | Single- or double-lepton triggers | |
| Lepton multiplicity | Exactly 2 leptons | |
| Lepton $p_T$ | $p_T > 25$ / 15 GeV | $p_T > 25$ / 15 (e) or 10 GeV ($\mu$) |
| Lepton $\eta$ | $\lvert\eta\rvert < 2.5$ (e) or 2.4 ($\mu$) | |
| $\tau_h$ multiplicity | No $\tau_h$ (loose WP) | $\geq 1$ $\tau_h$ (loose WP) and $< 2\tau_h$ (medium WP) |
| $\tau_h$ $p_T$ | — | $p_T > 20$ GeV |
| $\tau_h$ $\eta$ | — | $\lvert\eta\rvert < 2.3$ |
| Charge requirements | 2 same-sign leptons and charge quality requirements $\sum_{\ell,\tau_h} q = \pm 1$ | |
| Jet multiplicity | $\geq 4$ jets | $\geq 3$ jets |
| b tagging requirements | $\geq 1$ medium b-tagged jet or $\geq 2$ loose b-tagged jets | |
| Missing transverse momentum | $L_D > 30$ GeV ** | |
| Dilepton mass | $m_{\ell\ell} > 12$ GeV * and $\lvert m_{ee} - m_Z\rvert > 10$ GeV ** | |

| Selection | $1\ell$ + $2\tau_h$ | $2\ell$ + $2\tau_h$ |
|---|---|---|
| Targeted $t\bar{t}H$ decays | $t \to b\ell\nu$, $t \to bqq$, $H \to \tau\tau \to \tau_h\tau_h + \nu$'s | $t \to b\ell\nu$, $t \to b\ell\nu$, $H \to \tau\tau \to \tau_h\tau_h + \nu$'s |
| Trigger | Single-lepton or lepton+$\tau_h$ triggers | Single-, double-lepton triggers |
| Lepton multiplicity | Exactly 1 lepton | $\geq 2$ leptons |
| Lepton $p_T$ | $p_T > 25$ (e) or 20 GeV ($\mu$) | $p_T > 25$ / 15 (e) or 10 GeV ($\mu$) |
| Lepton $\eta$ | $\lvert\eta\rvert < 2.1$ | $\lvert\eta\rvert < 2.5$ (e) or 2.4 ($\mu$) |
| $\tau_h$ multiplicity | $\geq 2$ $\tau_h$ (medium WP) | |
| $\tau_h$ $p_T$ | $p_T > 30$ / 20 GeV | $p_T > 20$ GeV |
| $\tau_h$ $\eta$ | $\lvert\eta\rvert < 2.3$ | $\lvert\eta\rvert < 2.3$ |
| Charge requirements | $\sum_{\tau_h} q = 0$ | $\sum_{\ell,\tau_h} q = 0$ |
| Jet multiplicity | $\geq 3$ jets | $\geq 2$ jets |
| b tagging requirements | $\geq 1$ medium b-tagged jet or $\geq 2$ loose b-tagged jets | |
| Missing transverse momentum | — | No requirement if $N_{jet} \geq 4$ |
| | — | $L_D > 45$ GeV † |
| | — | $L_D > 30$ GeV otherwise |
| Dilepton mass | $m_{\ell\ell} > 12$ GeV * | |

* Applied on all pairs of leptons that pass loose selection.

** If both leptons are electrons.

† If the event contains a same-flavor opposite-sign lepton pair and $N_{jet} \leq 3$.

Table 5.9: Event selections applied in the $3\ell$, $3\ell + 1\tau_h$ and $4\ell$ categories.

| Selection | $3\ell$ | $3\ell + 1\tau_h$ |
|---|---|---|
| Targeted $t\bar{t}H$ decays | $t \to b\ell\nu$, $t \to b\ell\nu$, $H \to WW \to \ell\nu qq$<br>$t \to b\ell\nu$, $t \to bqq$, $H \to WW \to \ell\nu\ell\nu$ | $t \to b\ell\nu$, $t \to b\ell\nu$,<br>$H \to \tau\tau \to \ell\tau_h + \nu$'s |
| Trigger | Single-, double- or triple-lepton triggers | |
| Lepton multiplicity | Exactly 3 leptons | |
| Lepton $p_T$ | $p_T > 25$ / 15 / 15 GeV | $p_T > 20$ / 10 / 10 GeV |
| Lepton $\eta$ | $\lvert\eta\rvert < 2.5$ (e) or 2.4 ($\mu$) | |
| $\tau_h$ | No $\tau_h$ (loose WP) | $\geq 1$ $\tau_h$ (loose WP) |
| $\tau_h$ $p_T$ | — | $p_T > 20$ GeV |
| $\tau_h$ $\eta$ | — | $\lvert\eta\rvert < 2.3$ |
| Charge requirements | $\sum_\ell q = \pm 1$ | $\sum_{\ell,\tau_h} q = 0$ |
| Jet multiplicity | $\geq$2 jets | |
| b tagging requirements | $\geq$1 medium b-tagged jet or $\geq$2 loose b-tagged jets | |
| Missing transverse momentum | No requirement if $N_{jet} \geq 4$<br>$L_D > 45$ GeV [†]<br>$L_D > 30$ GeV otherwise | |
| Dilepton mass | $m_{\ell\ell} > 12$ GeV [*] and $\lvert m_{\ell\ell} - m_Z\rvert > 10$ GeV [‡] | |
| Four-lepton mass | $m_{4\ell} > 140$ GeV [§] | — |

| Selection | $4\ell$ |
|---|---|
| Targeted $t\bar{t}H$ decays | $t \to b\ell\nu$, $t \to b\ell\nu$, $H \to WW \to \ell\nu\ell\nu$<br>$t \to b\ell\nu$, $t \to b\ell\nu$,<br>$H \to ZZ \to \ell\ell qq$ or $\ell\ell\nu\nu$ |
| Trigger | Single-, double- or triple-lepton triggers |
| Lepton multiplicity | $\geq 4$ leptons |
| Lepton $p_T$ | $p_T > 25$ / 15 / 15 / 10 GeV |
| Lepton $\eta$ | $\lvert\eta\rvert < 2.5$ (e) or 2.4 ($\mu$) |
| Charge requirements | — |
| Jet multiplicity | $\geq$2 jets |
| b tagging requirements | $\geq$1 medium b-tagged jet or $\geq$2 loose b-tagged jets |
| Missing transverse momentum | — |
| Dilepton mass | $m_{\ell\ell} > 12$ GeV [*] and $\lvert m_{\ell\ell} - m_Z\rvert > 10$ GeV [‡] |
| Four-lepton mass | $m_{4\ell} > 140$ GeV [§] |

[*] Applied on all pairs of leptons that pass loose selection.
[†] If the event contains a same-flavor opposite-sign (SFOS) lepton pair and $N_{jet} \leq 3$.
[‡] Applied to all SFOS lepton pairs.
[§] Applied only if the event contains 2 SFOS lepton pairs.

### 5.3.3 Kinematics

The selection criteria on $p_T$ and $\eta$ are mostly driven by the trigger and object identification requirements. In categories without $\tau_h$, i.e. $2\ell$ss, $3\ell$, and $4\ell$, the highest and the second highest lepton $p_T$ are required to be greater than 25 and 15 GeV, respectively. The third highest lepton $p_T$, if applicable, needs to be greater than 15 GeV, and the $p_T$ threshold for the fourth lepton is 10 GeV. In case of $2\ell ss + 1\tau_h$ and $2\ell + 2\tau_h$, the leading lepton $p_T$ is required to be greater than 25 GeV. The trailing lepton needs to have a $p_T$ no less than 15 GeV if it is an electron, or 10 GeV if it is a muon. In the $3\ell + 1\tau_h$ category, the lepton with the highest $p_T$ is required to have a $p_T$ greater than 20 GeV, and both the second and third highest $p_T$ leptons need to satisfy the condition $p_T > 10$ GeV. In case of $1\ell + 2\tau_h$, the lepton $p_T$ is required to be greater than 20 GeV if it is a muon, and the threshold is increased to 25 GeV if it is an electron. In all categories except $1\ell + 2\tau_h$, any muons are required to be within the $\eta$ range $|\eta| < 2.4$. The $\eta$ requirement is extended to $|\eta| < 2.5$ in case of electrons. For the $1\ell + 2\tau_h$ category, the lepton, either an electron or a muon, is required to satisfy $|\eta| < 2.1$ because of the lepton+$\tau_h$ cross triggers used in this category. The $p_T$ of selected hadronic $\tau$ leptons in all categories are required to be greater than 20 GeV, except that the threshold for the leading $\tau_h$ $p_T$ is increased to 30 GeV in the $1\ell + 2\tau_h$ category. The $\eta$ for $\tau_h$'s in all categories are required to satisfy $|\eta| < 2.3$.

### 5.3.4 Charge requirements

Charge requirements are applied on leptons and $\tau_h$'s mainly to suppress the $t\bar{t}$ + jets background process. In $2\ell$ss and $2\ell ss + 1\tau_h$ categories, as the names sug-

gest, the two leptons are required to have the same sign of charge. Additional criteria are required on the quality of leptons' charge measurements in order to reduce backgrounds arising from misidentification of the lepton charge. For muons, this means the relative uncertainty of the global muon track $p_{\mathrm{T}}$ measurement needs to be less than 20%. For electrons, the charge measurement based on KF and GSF tracking algorithms (c.f. Section 4.2 and 4.5) as well as on hits reconstructed in the silicon pixel and strip trackers are required to be consistent.In case of $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$, since one of the leptons is expected to come from the leptonic $\tau$ decay, the charge of $\tau_{\mathrm{h}}$ is required to be of opposite sign with the leptons.

In the $1\ell + 2\tau_{\mathrm{h}}$ category, charges of the two $\tau_{\mathrm{h}}$'s are required to be of opposite sign. The charge sum of all leptons and $\tau_{\mathrm{h}}$'s in $2\ell + 2\tau_{\mathrm{h}}$ and $3\ell + 1\tau_{\mathrm{h}}$ categories are required to be zero. In case of the $3\ell$ category, the sum of the three leptons' charge must be either $+1$ or $-1$. No charge requirement is applied in the $4\ell$ category.

### 5.3.5 Jet and b tagging requirements

The presence of multiple jets are expected in all event categories. Events selected in the $2\ell\mathrm{ss}$ category are required to have at least four jets. In case of $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$ and $1\ell + 2\tau_{\mathrm{h}}$, at least three jets are required. The requirement on the jet multiplicity is further reduced to at least two in rest of the event categories containing more leptons and hadronic taus. Events in all categories also have to contain at least one medium b tags or two loose b tags based on the DeepCSV b tagging algorithm (c.f. Section 4.7). The requirement on the presence of jets originated from b quarks suppresses efficiently those backgrounds from electroweak processes

in which no top or anti-top quarks are produced.

### 5.3.6 Missing transverse momentum

The linear discriminator $L_{\mathrm{D}}$, defined in Eq. 5.3, is used to select events with missing transverse momentum. The requirement on the presence of $p_{\mathrm{T}}^{\mathrm{miss}}$ helps to reduce the Z+jets background events, especially when one of the leptons (most likely electrons) from the Z boson decay has its charge mismeasured, so that the charge requirements would not help. In $2\ell$ss and $2\ell$ss + $1\tau_{\mathrm{h}}$ categories, $L_{\mathrm{D}}$ is required to be greater than 30 GeV only if both the leptons are electrons, since the misidentification rate of electron charges is much higher than that of muons. In case of $2\ell$ + $2\tau_{\mathrm{h}}$, $3\ell$, and $3\ell$ + $1\tau_{\mathrm{h}}$ categories, no $L_{\mathrm{D}}$ requirement is applied if the event contains more than three jets. Otherwise, $L_{\mathrm{D}}$ is required to be greater than 45 GeV if the event contains a pair of same-flavor opposite-sign (SFOS) leptons passing loose selection criteria. The threshold on $L_{\mathrm{D}}$ is lowered to 30 GeV if no such lepton pair is found. There is no $L_{\mathrm{D}}$ requirement in the categories of $1\ell+2\tau_{\mathrm{h}}$ and $4\ell$.

In all event categories, a set of $p_{\mathrm{T}}^{\mathrm{miss}}$ filters are applied to remove events with spurious $p_{\mathrm{T}}^{\mathrm{miss}}$ due to instrumental effects such as calorimeter noise, beam halo, badly reconstructed vertices, etc. The $p_{\mathrm{T}}^{\mathrm{miss}}$ filters applied to both the recorded and simulated data are listed in Table 5.10. Details about the $p_{\mathrm{T}}^{\mathrm{miss}}$ filters can be found in Ref. [65].

Table 5.10: $p_{\mathrm{T}}^{\mathrm{miss}}$ filters applied to collision data and to simulated events. A hyphen (−) indicates that the filter is not applied.

| Filter name | Applied to data | Applied to simulation |
|---|---|---|
| Flag_goodVertices | ✓ | ✓ |
| Flag_globalTightHalo2016Filter | ✓ | ✓ |
| Flag_HBHENoiseFilter | ✓ | ✓ |
| Flag_HBHENoiseIsoFilter | ✓ | ✓ |
| Flag_EcalDeadCellTriggerPrimitiveFilter | ✓ | ✓ |
| Flag_BadPFMuonFilter | ✓ | ✓ |
| Flag_BadChargedCandidateFilter | ✓ | ✓ |
| Flag_eeBadScFilter | ✓ | − |
| Flag_ecalBadCalibFilter | ✓ | ✓ |

### 5.3.7 Lepton mass veto

Events in any categories containing a pair of leptons passing loose selection criteria with an invariant mass below 12 GeV are excluded in this analysis, because these events are not well modeled in the MC simulated samples.

In order to further reject background events involving Z bosons, an additional requirement on the invariant mass of pairs of loose leptons is applied to all event categories except $1\ell + 2\tau_{\mathrm{h}}$ and $2\ell + 2\tau_{\mathrm{h}}$. Events would be vetoed if they contain a pair of these leptons with the invariant mass less than 10 GeV away from the Z boson mass ($\mathrm{m_Z}$ = 91.2 GeV). In $2\ell$ss and $2\ell$ss + $1\tau_{\mathrm{h}}$ categories, the Z-mass veto is applied only if the pair of leptons are electrons, since the muon charge misidentification rate is negligibly small and the same-sign requirement is sufficient to suppress $\mathrm{Z} \to \mu\mu$ decays. In case of $3\ell$, $3\ell + 1\tau_{\mathrm{h}}$, and $4\ell$ categories, the Z-mass veto is applied only if the lepton pair are SFOS.

For $3\ell$ and $4\ell$ categories, events are vetoed if they contains two SFOS pairs of loose leptons and the invariant mass of these four leptons is less than 140 GeV. This four-lepton mass veto is to avoid the potential overlap with other analyses

targeting $H \to ZZ^* \to 4\ell$ decay channels.

## 5.4 Modeling of signal and background processes

Data events selected based on the criteria described in Section 5.3 are to be compared with the predictions under either background-only or background-plus-signal hypothesis. The estimation of $t\bar{t}H$ events in the signal region (SR) of each event category is based on the MC simulation. Non-$t\bar{t}H$ events could be selected in a SR if their final states contain the same number of genuine prompt leptons and/or $\tau_h$'s as required by an event category. Such processes are categorized as "irreducible" backgrounds. The other type of background, referred to as "reducible" backgrounds, arises from mismeasurement of leptons or $\tau_h$'s, primarily in $t\bar{t}$ events. In such cases, a non-prompt muon or electron originating from the decay of a hadron passes the tight lepton selection and is misidentified as a prompt lepton, or a jet is misidentified as a prompt electron, muon, or $\tau_h$. In addition, in $2\ell$ss and $2\ell$ss + $\tau_h$ categories, backgrounds can also arise from events containing a pair of opposite-sign leptons, in which one lepton charge is mismeasured.

The irreducible and reducible backgrounds have comparable contributions in $2\ell$ss, $2\ell$ss + $1\tau_h$, $3\ell$, and $3\ell$ + $1\tau_h$ categories. The reducible backgrounds are dominant in the categories requiring two $\tau_h$'s due to the relatively higher probability of misidentifying jets as $\tau_h$'s. In the $4\ell$ category, backgrounds are mostly of irreducible type.

### 5.4.1 Background from misidentified leptons and $\tau_h$'s

The reducible background due to misidentification of prompt leptons and $\tau_h$'s, referred to hereafter as "misidentified" or "non-prompt" background, is estimated from data using the fake factor (FF) method. To apply the fake factor method, an "application region" (AR) is defined in each event category. The selection criteria for each AR are exactly the same as its corresponding SR except that the lepton and $\tau_h$ selection requirements are relaxed to "fakeable" instead of "tight" (c.f. Table 5.4, 5.3, and 5.6). Events that would be selected in the SR are excluded from the AR in order to ensure the SR and AR are mutually exclusive. The contribution of misidentified background in the SR is then estimated by extrapolating from the AR based on fake factors.

The fake factor, denoted as $f_i$, is defined as the probability of a misidentified electron, muon, or $\tau_h$ to pass the tight selection given that it satisfies the fakeable object selection criteria. The fake factors for electrons, muons, and $\tau_h$'s are measured in dedicated control regions referred to as "determination regions" (DR). The DRs are chosen so that the selected fakeable object is unlikely to be a genuine or prompt lepton or $\tau_h$. The DRs for measuring electron and muon fake factors are dominated by multijet events. For $\tau_h$ fake factors, the DR is comprised of $t\bar{t}$ + jets events. The fake factors are parameterized as functions of $p_T$ (cone-$p_T$ in case of leptons) and $\eta$. More details on lepton and $\tau_h$ fake factor measurements are discussed in Appendix A.1 and A.2.

The extrapolation from AR to SR in each event category for misidentified backgrounds is carried out by applying weights to events in the AR. The event weight that is used to convert yields in the AR to the SR is computed based on fake factors and the number of objects that fails the tight selections in the AR.

For event categories with two objects, in a given $p_{\mathrm{T}}$ (cone-$p_{\mathrm{T}}$ in case of leptons) and $\eta$ bin, the number of events in the SR can be denoted as $N_{pp}$ and the number of events in the AR is $N_{pf}+N_{fp}+N_{ff}$. The subscript $p$ and $f$ stand for passing and failing tight object selections, respectively. The order of the subscript matches to that of selected objects in descending $p_{\mathrm{T}}$ order. These four numbers can be expressed in terms of the numbers of events in which both objects are prompt ($N_{11}$), one of the objects is non-prompt/fake ($N_{10}$ and $N_{01}$), and both are non-prompt/fake ($N_{00}$), as shown in Eq. 5.4.

$$\begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} = \begin{pmatrix} \epsilon_1\epsilon_2 & \epsilon_1 f_2 & f_1\epsilon_2 & f_1 f_2 \\ \epsilon_1(1-\epsilon_2) & \epsilon_1(1-f_2) & f_1(1-\epsilon_2) & f_1(1-f_2) \\ (1-\epsilon_1)\epsilon_2 & (1-\epsilon_1)f_2 & (1-f_1)\epsilon_2 & (1-f_1)f_2 \\ (1-\epsilon_1)(1-\epsilon_2) & (1-\epsilon_1)(1-f_2) & (1-f_1)(1-\epsilon_2) & (1-f_1)(1-f_2) \end{pmatrix} \begin{pmatrix} N_{11} \\ N_{10} \\ N_{01} \\ N_{00} \end{pmatrix} \tag{5.4}$$

The efficiency of the a prompt object passing the tight selection is denoted as $\epsilon_i$, and $f_i$ is the fake factor for the misidentified object. The index $i$ indicates the efficiency or the fake factor is for the $i$-th object. The above linear equation can be simplified assuming the probability for a prompt object to fail the tight selection is very small so that the contribution to the AR ($N_{pf}$, $N_{fp}$, $N_{ff}$) from prompt objects is negligible compared to that from non-prompt/fake objects. Therefore, Eq. 5.4 can be simplified to

$$\begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} = \begin{pmatrix} \epsilon_1\epsilon_2 & \epsilon_1 f_2 & f_1\epsilon_2 & f_1 f_2 \\ 0 & \epsilon_1(1-f_2) & 0 & f_1(1-f_2) \\ 0 & 0 & (1-f_1)\epsilon_2 & (1-f_1)f_2 \\ 0 & 0 & 0 & (1-f_1)(1-f_2) \end{pmatrix} \begin{pmatrix} N_{11} \\ N_{10} \\ N_{01} \\ N_{00} \end{pmatrix}. \tag{5.5}$$

Based on the equation above, $N_{00}$, $N_{10}$, and $N_{01}$ can be expressed in terms of $N_{pf}$, $N_{fp}$, and $N_{ff}$:

$$N_{00} = \frac{N_{ff}}{(1-f_1)(1-f_2)} \tag{5.6}$$

$$N_{10} = \frac{N_{pf} - \frac{f_1}{1-f_1} N_{ff}}{\epsilon_1 (1 - f_2)} \tag{5.7}$$

$$N_{01} = \frac{N_{fp} - \frac{f_2}{1-f_2} N_{ff}}{\epsilon_2 (1 - f_1)} \tag{5.8}$$

The number of events with both objects passing the tight selection due to non-prompt/fake objects is

$$\begin{aligned} N_{pp}^{bkg} &= \epsilon_1 f_2 N_{10} + \epsilon_2 f_1 N_{01} + f_1 f_2 N_{00} \\ &= \epsilon_1 f_2 \frac{N_{pf} - \frac{f_1}{1-f_1} N_{ff}}{\epsilon_1 (1 - f_2)} + \epsilon_2 f_1 \frac{N_{fp} - \frac{f_2}{1-f_2} N_{ff}}{\epsilon_2 (1 - f_1)} + f_1 f_2 \frac{N_{ff}}{(1 - f_1)(1 - f_2)} \\ &= \frac{f_1}{1 - f_1} N_{pf} + \frac{f_2}{1 - f_2} N_{pf} - \frac{f_1 f_2}{(1 - f_1)(1 - f_2)} N_{ff} \end{aligned} \tag{5.9}$$

The weight applied to an event in the AR for extrapolation in case of event categories with two objects is therefore:

$$w_2 = \begin{cases} \frac{f}{1-f} & \text{if one object fails the tight selection} \\ \frac{-f_1 f_2}{(1-f_1)(1-f_2)} & \text{if both object fail the tight selection} \end{cases} \tag{5.10}$$

Similar arguments can be made for event categories with three or more objects. The generic event weight to estimate the non-prompt background in the SR from events in the AR can be expressed as:

$$w_{\mathrm{FF}} = (-1)^{n_f+1} \prod_{i=1}^{n_p+n_f} \begin{cases} 1 & \text{if object } i \text{ passes the tight selection} \\ \frac{f_i}{1-f_i} & \text{if object } i \text{ fails the tight selection} \end{cases} \tag{5.11}$$

$n_p$ and $n_f$ are numbers of objects that pass tight selections and that fail tight selections in an event, respectively. By applying event weights defined in Eq. 5.11 to data sample selected in the AR, an unbiased estimation of the non-prompt background in the SR can be obtained. The estimation is further improved by applying a correction to account for the contamination of irreducible background events selected in the AR. The correction is obtained from MC simula-

tion by carrying out the same extrapolation to the simulated irreducible background events in the AR. The correction is applied by subtracting this extrapolation from that obtained from data in the AR.

The FF method is applied as described above in $2\ell$ss, $3\ell$, $4\ell$, $1\ell + 2\tau_h$, and $2\ell + 2\tau_h$ categories. Since the misidentified background is estimated from data with the FF method, selected leptons or $\tau_h$'s of all MC simulated events in the SR are required to be compatible with generator level prompt leptons or genuine $\tau_h$ with $\Delta R < 0.2$. This generator level matching requirement reduces contamination from non-prompt objects in MC samples and avoids over-estimating the misidentified background.

In $2\ell$ss $+ 1\tau_h$ and $3\ell + 1\tau_h$ categories, only background events due to misidentification of electrons and/or muons are estimated using the FF method. Therefore, in these two categories, only the electron and muon selection criteria are relaxed to "fakeable" in the AR. The reweighting is not applied to the $\tau_h$ candidate. The background due to misidentifying jets as $\tau_h$'s in these two categories are estimated using MC simulated samples by matching the selected $\tau_h$ to the generator level hadronic $\tau$ lepton. For each simulated signal and background process, events are grouped based on MC matching results into genuine or fake $\tau_h$ categories. This modified FF method allows $t\bar{t}H$ events that contains misidentified $\tau_h$ are still treated as signals. Such events, which are mainly from $H \to WW$ decays in the $t\bar{t}H$ process, amount to about 30% of the total signal events in these categories. The analysis sensitivity is therefore optimized for the inclusive $t\bar{t}H$ production rather than only the targeted final states in these two categories.

Closure tests are carried out to validate estimations of misidentified backgrounds using the FF method. MC samples of $t\bar{t}$ + jets and QCD processes are

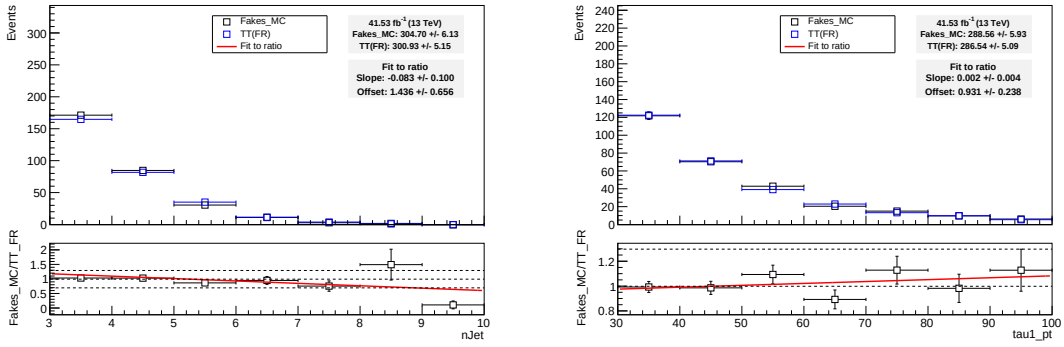


Figure 5.1: Closure tests using simulated $t\bar{t}$ + jets events in the $1\ell + 2\tau_h$ category. Distributions of the number of jets (left) and leading $\tau_h$ candidate $p_T$ (right) extrapolated from the AR using the FF method ("TT(FR)", in blue) are compared to those obtained directly from the SR ("Fakes_MC", in black).

used for the closure test. The simulated events selected in the AR of each event category are extrapolated to the SR as described above. The extrapolation is then compared to the events directly selected in the SR. An example of the closure test in the $1\ell + 2\tau_h$ category is shown in Fig. 5.1, in which distributions of the number of jets as well as $p_T$ of the leading $\tau_h$ candidate are plotted. Good agreement between extrapolations from AR and events selected in the SR are observed. The residual discrepancies are accounted for as systematic uncertainties.

### 5.4.2 Background from lepton charge misidentification

In $2\ell$ss and $2\ell$ss + $1\tau_h$ categories, a $t\bar{t}$ + jets event, in which two prompt opposite-sign leptons can be produced from dileptonic decay of top quark pair, could be mistakenly selected into the SR if the charge of either lepton is mismeasured. This background, referred to as "charge mismeasured" or "sign-flip" background, is estimated from data control regions similarly as the estimation

of misidentified background. Application regions are defined with the same selection criteria as the signal regions of $2\ell$ss and $2\ell$ss + $1\tau_h$ categories, except charges of the two leptons are required to be different. Events in the sign-flip AR are reweighted to estimate the contribution in the SR due to charge mismeasurement. The event weight applied in the extrapolation is determined based on the probability to misidentify lepton charges. For events in the $2\ell$ss sign-flip AR, either lepton charge could be mismeasured. The event weight for extrapolation, i.e. the probability of such event to pass the SR event selection, is therefore the sum of the charge misidentification rates of the two leptons. In case of events in the $2\ell$ss+$1\tau_h$ sign-flip AR, since the SR selection requires the $\tau_h$ and the leptons to be of opposite charge, only one of the two leptons could potentially contribute to the sign-flip backgrounds if its charge is misidentified. The event weight for extrapolation is therefore set to the charge misidentification probability of the lepton that has the same charge as the $\tau_h$ in the AR.

The charge misidentification rate is measured separately for electrons and muons in Drell-Yan events ($Z/\gamma^* \rightarrow ee$, $Z/\gamma^* \rightarrow \mu\mu$) in bins of $p_T$ and $\eta$. For an electron, the charge misidentification rate varies roughly from 0.01% to 0.15% depending on its $p_T$ and $\eta$. For a muon, the charge misidentification rate is found to be negligible. More information on the measurement of lepton charge misidentification rate can be found in Appendix A.3.

### 5.4.3 Background from photon conversion

An additional small reducible background, mostly relevant in $2\ell$ss and $2\ell$ss+$1\tau_h$ categories, comes from $t\bar{t}$ production in association with either a real or a virtual

photon ($t\bar{t}\gamma/\gamma^*$). Such background events would be selected in the SR if the photon converts to an electron-positron pair, in which one of them carries most of the energy while the other is too soft to be reconstructed. This background process is suppressed effectively by the photon conversion veto and no missing pixel hits of the electron track criteria required in fakeable and tight electron selections (c.f. Table 5.4). In addition, the $m_{\ell\ell} > 12$ GeV requirement in event selections (c.f. Section 5.3.7) also helps to reduce $t\bar{t}$ events associated with a virtual photon, since the electron or muon pair produced from the virtual photon is usually of low mass. The remaining small background due to photon conversions are modeled using MC samples.

### 5.4.4 Irreducible backgrounds

The main irreducible background process is the production of a pair of top quarks associated with a W or Z vector boson ($t\bar{t}W$ or $t\bar{t}Z$). They are sometimes referred to together as $t\bar{t}V$. The production of a top quark pair with two W bosons ($t\bar{t}WW$) is considered in this analysis as well. The production of a vector boson pair (WZ, ZZ, WW) in association with jets also has considerable contributions especially in the event categories with three or more leptons. Due to the small cross section of the $t\bar{t}H$ signal process, backgrounds from relatively rare standard model processes are also considered. These rare background processes include the production of three vector bosons (WWW, WWZ, WZZ, ZZZ), the single top production associated with a Z boson (tZq), the production of two W bosons with same sign (WpWp), WW double scattering, and the production of four top quarks ($t\bar{t}t\bar{t}$). All irreducible backgrounds are modeled based on MC simulations with their corresponding cross sections set to the SM predic-

tions. Total yields of these simulated processes are normalized according to the amount of recorded data used in the analysis, except the $t\bar{t}W$ and $t\bar{t}Z$ processes. These two background processes are not fixed to the SM expectations but determined from the fit during the signal extraction. Signal extraction strategies are discussed in more details in Section 5.5.

In this version of the analysis, the production of a single top quark associated with a Higgs boson (tH) is treated as a background. The tH MC sample is generated assuming destructive interference between the two underlying processes based on the standard model prediction.

### 5.4.5 Data-to-MC corrections

Potential differences in e.g. detector response and algorithm performance between collision data and simulations need to be accounted for to improve modeling of signal and background processes using MC simulations. In addition to corrections to the energy scale and resolution of reconstructed physics objects as discussed in Chapter 4, each simulated event is reweighted by an event weight, which is calculated per event as a product of several scale factors that account for different effects described below.

#### Pileup

To mimic the pileup condition at the LHC, simulated hard scattering processes in MC samples are mixed with minimum bias events. The number of the minimum bias events is drawn randomly for each event from an expected pileup

distribution. The expected pileup distribution, however, does not always match exactly the pileup distribution measured in the recorded data. The isolation criteria required in the physics object selections as described in Section 5.2 are designed to mitigate the pileup effect on the analysis. To further account for the difference in pileup between data and simulation, a scale factor is determined from the ratio of the measured pileup distribution to that in MC samples. Each simulated event is then reweighted using this scale factor.

**Trigger efficiency**

The high-level triggers listed in Table 5.7 are applied in both data and simulated events. Lepton trigger efficiencies are determined in both data and simulations. To measure trigger efficiencies in data events, a trigger requiring missing transverse momentum above a certain threshold is used to record data for this measurement. The $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger is uncorrelated with the lepton triggers used in all analysis categories, therefore no trigger bias is introduced in the measurement. Trigger efficiencies in data and simulation are found to be generally consistent. The small discrepancy is corrected by trigger scale factors, which are calculated from the ratio of the trigger efficiency in data to that in the simulation. Trigger efficiencies, thus trigger scale factors, are parameterized as a function of lepton flavor, multiplicity, and leading lepton $p_{\mathrm{T}}$ in the event. The trigger scale factors and their uncertainties are summarized in Table 5.11 for event categories using single-, double-, and/or triple-lepton triggers. In the $1\ell + 2\tau_{\mathrm{h}}$ category, an additional complication is introduced by the combination of single-lepton triggers and lepton-$\tau_{\mathrm{h}}$ cross triggers. Efficiencies to pass the single lepton triggers ($\epsilon_L$), the lepton leg of the lepton-$\tau_{\mathrm{h}}$ cross triggers ($\epsilon_\ell$), and the $\tau_{\mathrm{h}}$ leg of the

Table 5.11: Trigger efficiency scale factors applied to simulated events for $2\ell$ss$(+1\tau_{\mathrm{h}})$, $3\ell(+1\tau_{\mathrm{h}})$, and $4\ell$ categories.

| Lepton type | Leading lepton $p_{\mathrm{T}}$ | Scale factor |
|---|---|---|
| $2\,\mu$ | $p_{\mathrm{T}} < 35$ GeV | $0.972 \pm 0.006$ |
| | $p_{\mathrm{T}} \geq 35$ GeV | $0.994 \pm 0.001$ |
| $\mathrm{e} + \mu$ | $p_{\mathrm{T}} < 35$ GeV | $0.952 \pm 0.008$ |
| | $35 \leq p_{\mathrm{T}} < 50$ GeV | $0.983 \pm 0.003$ |
| | $p_{\mathrm{T}} \geq 50$ GeV | $1.000 \pm 0.001$ |
| $2\,\mathrm{e}$ | $p_{\mathrm{T}} < 30$ GeV | $0.937 \pm 0.027$ |
| | $p_{\mathrm{T}} \geq 30$ GeV | $0.991 \pm 0.002$ |
| $\geq 3\,\ell$ | – | $1.000 \pm 0.050$ |

cross triggers ($\epsilon_\tau$) are measured as a function of $p_{\mathrm{T}}$ and $\eta$ in $\mathrm{Z}/\gamma^* \rightarrow$ ee, $\mu\mu$, $\tau\tau$ events using the tag-and-probe [83] method. The efficiency of a $1\ell + 2\tau_{\mathrm{h}}$ event to pass the combination of these triggers can be expressed in terms of $\epsilon_L$, $\epsilon_\ell$, $\epsilon_{\tau 1}$, and $\epsilon_{\tau 2}$, where $\tau 1$ and $\tau 2$ denote the $\tau_{\mathrm{h}}$ with the highest and second highest $p_{\mathrm{T}}$, respectively. The efficiency $\epsilon$, expression of which is given in Eq. 5.12, depends on which trigger the event fires. The equation takes into account that either or both $\tau_{\mathrm{h}}$'s could fire the $\tau_{\mathrm{h}}$ leg of the cross trigger and that the efficiency of the standalone lepton trigger is not necessarily always lower than the lepton leg of the cross trigger.

$$\epsilon = \begin{cases} \epsilon_L - \min(\epsilon_L, \epsilon_\ell) \times (1 - (1 - \epsilon_{\tau 1})(1 - \epsilon_{\tau 2})) & \text{if only the single lepton trigger fires} \\ (\epsilon_\ell - \epsilon_L) \times (1 - (1 - \epsilon_{\tau 1})(1 - \epsilon_{\tau 2})) & \text{if only the cross-trigger fires} \\ \min(\epsilon_L, \epsilon_\ell) \times (1 - (1 - \epsilon_{\tau 1})(1 - \epsilon_{\tau 2})) & \text{if both triggers fire} \end{cases} \tag{5.12}$$

The trigger efficiencies in both data and simulation can be computed separately following the above equation, and the trigger scale factor in the $1\ell + 2\tau_{\mathrm{h}}$ category is given by their ratio.

### Lepton selection efficiency

Muon and electron selection efficiencies are measured separately in data and MC simulation using the tag-and-probe method as a function of $p_T$ and $\eta$ in $t\bar{t}$ + jets events. The total selection efficiency for a signal lepton is factorized into two components: the efficiency of a reconstructed lepton to pass the loose selection criteria and the efficiency of a loose lepton to pass the tight selection criteria. In case of the latter, the efficiency is measured separately with and without the lepton charge quality requirement for event selections in $2\ell ss(+1\tau_h)$ and other categories, respectively. The measured efficiency of loose muons or electrons to pass the tight selection with charge quality requirement as a function of $p_T$ in different $\eta$ bins are shown in Fig. 5.2. The ratio of the overall selection efficiency in data to that in MC simulation is calculated for each lepton. The scale factor for reweighting each simulated event is then obtained as the product of these per-lepton ratios.

### Selection efficiency of $\tau_h$

The efficiency of a $\tau_h$ candidate to pass the tight selection criteria are measured in $Z/\gamma^* \to \tau\tau$ and $t\bar{t} \to b\nu\mu\bar{b}\nu\tau$ events using the tag-and-probe method. Details of the $\tau_h$ selection efficiency measurement can be found in Ref. [63]. The ratio of the efficiency measured in data and in simulation is found to be different for different $\tau_h$ MVA ID working points. The ratio or scale factor, given in Table 5.12, is applied to each $\tau_h$ candidate if it matches to a genuine hadronic $\tau$ decay on generator level.

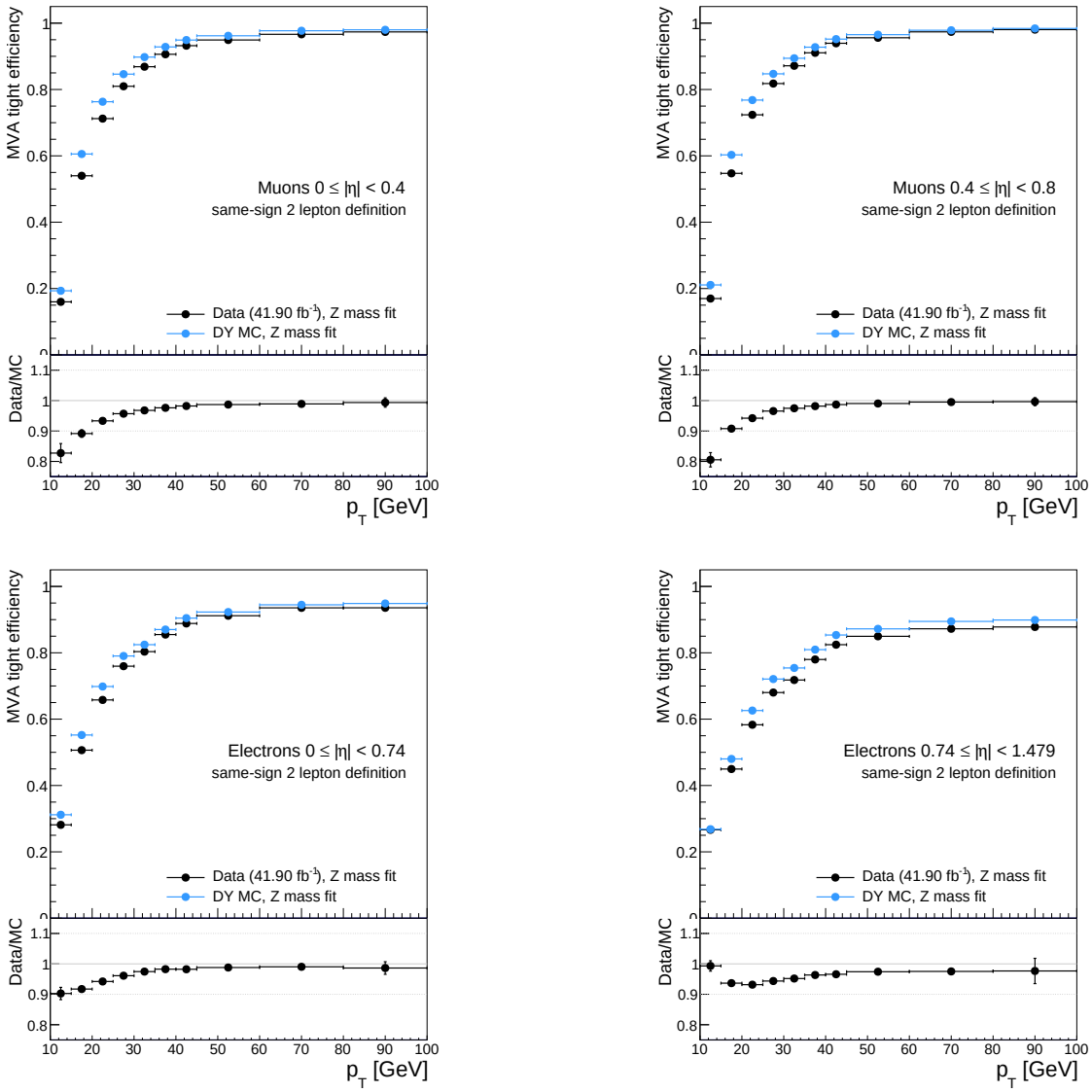


Figure 5.2: Lepton selection efficiencies as functions of $p_{\mathrm{T}}$ of loose muon and electron to pass the corresponding tight selection criteria. Muon efficiencies in $0 \leq |\eta| < 0.4$ (left) and $0.4 \leq |\eta| < 0.8$ (right) regions are shown on the top row. Electron efficiencies in $0 \leq< |\eta| < 0.74$ and $0.74 \leq |\eta| < 1.479$ regions are on the bottom row. Efficiencies measured in recorded data and MC simulated Drell-Yan (DY) samples are shown in black and blue markers, respectively. Data-to-MC ratios are also shown.

**b tagging**

The distribution of b tagging discriminants described in Section 4.7 may not be exactly the same in data and in simulation due to imperfect modeling of some

Table 5.12: Scale factors of $\tau_h$ selection efficiencies for Very-loose, Loose, and Medium $\tau_h$ MVA ID working point.

| $\tau_h$ MVA ID WP | Scale factor |
|---|---|
| Very-loose | 0.85 ± 0.05 |
| Loose | 0.89 ± 0.05 |
| Medium | 0.89 ± 0.05 |

input variables e.g. the impact parameter significance. In order to correct the potential distortion of b tagging discriminant distributions in simulated events, a per jet data-to-MC scale factor is obtained for each jet in an event. The per jet scale factor is the ratio of the b tagging efficiency in data to that in simulation, parameterized as a function of jet $p_T$, $\eta$, flavor, and the b tagging discriminant. Measurement of b tagging efficiencies in both data and simulation are detailed in Ref. [58]. The per event scale factor is computed by multiplying all per jet scale factors of selected jets in the event.

### 5.4.6 Background control regions

The modeling of background processes are validated in several control regions. These control regions are defined to be orthogonal to all signal regions and are designed to have dominant contributions from major background processes. Collision events in recorded data selected in these control regions are compared to estimations based on simulated samples and the FF method. Good agreements are observed in all control regions.

### $1\ell + 2\tau_h$ss **control region**

Event selection criteria for the $1\ell+2\tau_h$ss control region are the same as the $1\ell+2\tau_h$ SR, except the two $\tau_h$'s are required to be of the same sign instead of opposite sign. Figure 5.3 shows the total event yield and the distribution of di-$\tau_h$ invariant mass in this control region.

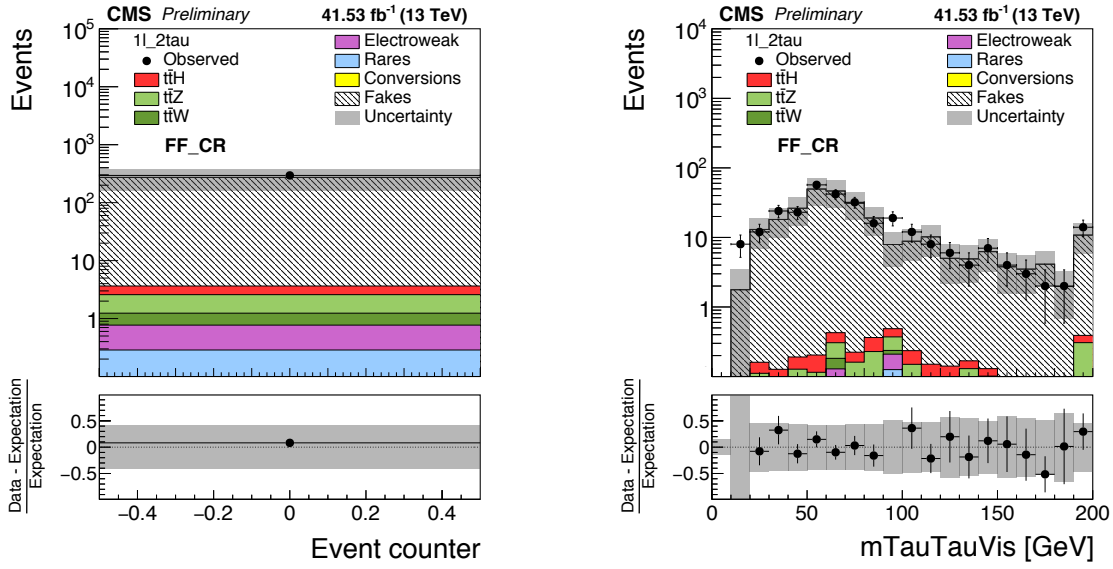


Figure 5.3: Event yields (left) and the distribution of di-$\tau_h$ invariant mass (right) in the $1\ell + 2\tau_h$ss control region.

### tt̄W **control region**

Event selection criteria for the tt̄W control region are based on those of the $2\ell$ss category. The only difference is that instead of requiring at least 4 jets, exactly 3 selected jets are required for events selected in this control region. The distribution of transverse missing momentum $E_T^{miss}$ as well as events yields in each $2\ell$ss sub-categories (c.f. Section 5.5.2) are shown in Fig. 5.4. The "bt" and "bl"

in the plot stand for "b-tight" and "b-loose" sub-categories, respectively. Event sub-categorization in $2\ell$ss and $3\ell$ categories is described in Section 5.5.2.

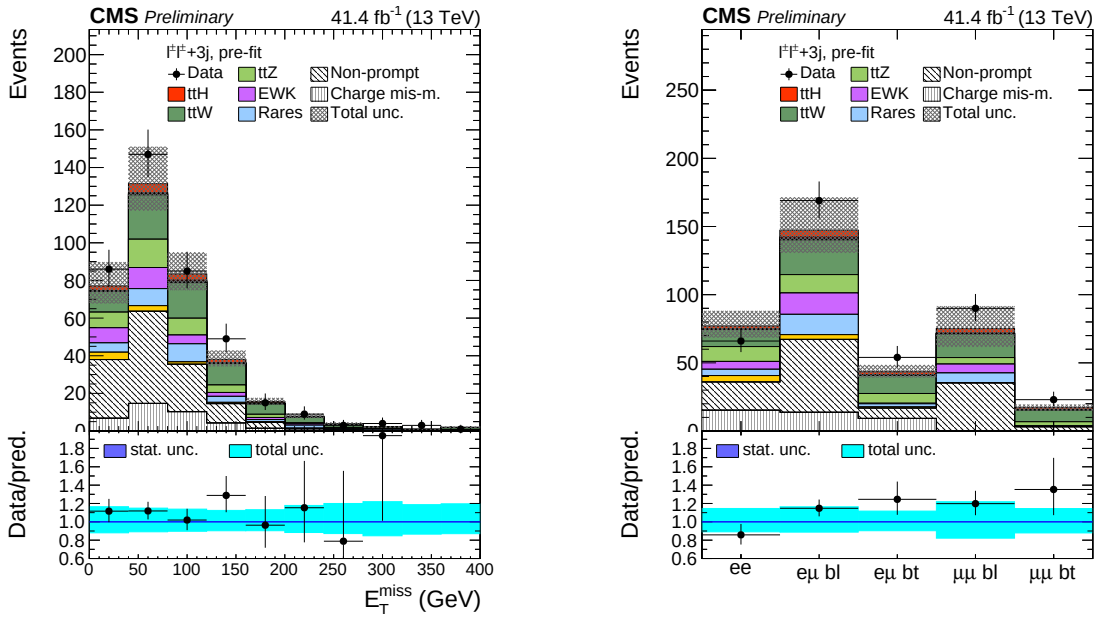


Figure 5.4: The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution (left) and event yields in $2\ell$ss sub-categories (right) in the $\mathrm{t\bar{t}W}$ control region.

### $\mathrm{t\bar{t}Z}$ control region

Event selection criteria defined in the $\mathrm{t\bar{t}Z}$ control region are the same as the $3\ell$ SR, except that events with a pair of SFOS leptons within 10 GeV of the Z boson mass are selected instead. The left plot in Fig. 5.5 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in this control region, and the plot on the right is a distribution of the dilepton mass that is closest to the Z boson mass in the event.

### WZ control region

Event selection criteria for the WZ control region are based on the $3\ell$ SR. The Z mass requirement in the $3\ell$ SR is inverted . In addition, the b tag requirement is

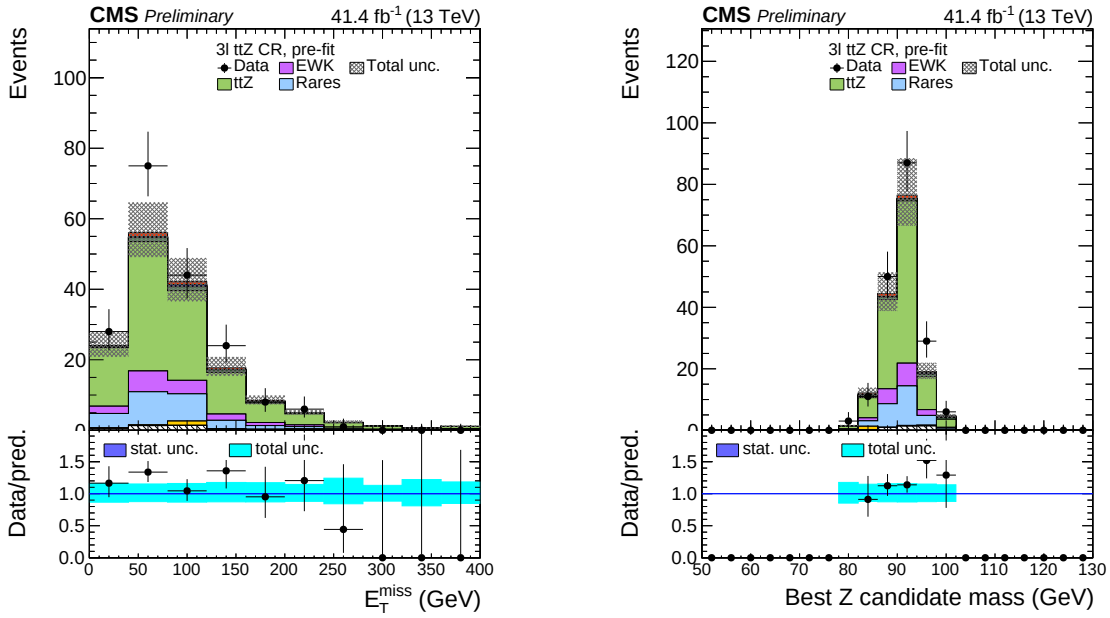


Figure 5.5: Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and dilepton mass that is closest to the mass of a Z boson in the $\mathrm{t\bar{t}Z}$ control region.

also inverted, i.e. events selected in the this control region are required to have at most one loose b-tagged jets. Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the number of selected jets in both data and the MC-based estimation are shown in Fig. 5.6.

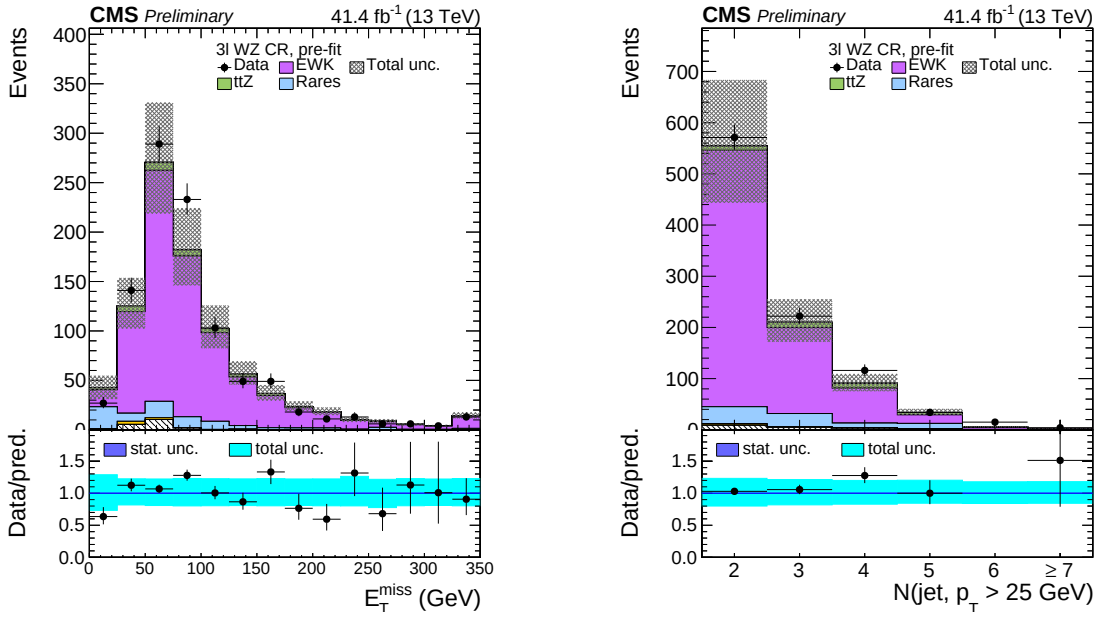


Figure 5.6: Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and the number of selected jets (right) in the WZ control region.

## 5.5 Signal extraction

The $t\bar{t}H$ cross section is measured in terms of the signal strength modifier $\mu_{t\bar{t}H} = \sigma/\sigma_{SM}$, where $\sigma$ and $\sigma_{SM}$ are the measured and SM predicted cross section, respectively. In order to enhance the sensitivity of the analysis, a discriminating observable is defined for each event category to maximize the separation between signal and background processes. The $t\bar{t}H$ signal strength is extracted from distributions of these discriminating observables by a binned maximum likelihood fit. The discriminating observable in each event category is described in more details in Section 5.5.1. Statistical treatments for extracting the signal strength is discussed in Section 5.5.2.

### 5.5.1 Discriminating observables

In each event category except $4\ell$, the discriminating observable is defined based on the output of boosted decision trees. Dedicated BDTs are trained in each event category on simulated events. In the event category with at least one hadronic taus (i.e. $1\ell + 2\tau_h$, $2\ell ss + 1\tau_h$, $3\ell + 1\tau_h$, or $2\ell + 2\tau_h$ categories), one BDT is trained using simulated $t\bar{t}H$ samples as the signal and the sum of major background processes ($t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}$ + jets) as the background. Each background process is weighted based on its relative event yields after event selections in each category. In the $2\ell$ or $3\ell$ category, two separate BDTs are trained to discriminate the $t\bar{t}H$ signal against $t\bar{t}W$ and $t\bar{t}Z$ backgrounds and against the $t\bar{t}$ + jets background. Due to the limited number of selected events in the $4\ell$ category, the signal extraction is based on total event yields in this category instead of a distribution.

Observables that can reflect the difference in kinematic features between signal and background processes are chosen as input variables for these BDTs. In different event categories, different sets of input variables are selected to optimize the BDT performance. The total number of input variables for each BDT is adjusted based on the number of simulated training events available in each category to avoid overtraining. BDT input variables used in each category are summarized in Table 5.13. Definitions of these variables are given below. The symbol $\ell$, $\tau$, and $j$ stand for a lepton (e or $\mu$), a hadronic $\tau$, and a jet, respectively. The subscript 1 and 2 indicates the object is the one with the highest ("leading") or the second highest ("trailing") $p_{\mathrm{T}}$ (cone-$p_{\mathrm{T}}$ in case of a lepton), respectively.

- cone-$p_{\mathrm{T}}^{\ell}$: cone-$p_{\mathrm{T}}$ of a lepton.
- max($|\eta^{\ell}|$): maximum $|\eta|$ among all leptons.
- $|\cos\theta^{\tau_1}|$: absolute value of cosine of the leading $\tau_{\mathrm{h}}$'s polar angle.
- $\Delta R(x, y)$: angular distance in $\eta$-$\phi$ coordinates between an object $x$ and an object $y$.
- $\Delta R_{\mathrm{ss}}(x, y)$: same as $\Delta R(x, y)$ except object $x$ and $y$ are required to be of the same sign.
- $\left\langle \Delta R_{\mathrm{jj}} \right\rangle$: average angular distance between selected jets.
- $\langle \Delta R_{\ell\tau} \rangle$: average angular distance between a lepton and a $\tau_{\mathrm{h}}$.
- min($\Delta R(\tau, j)$): minimum angular distance between any $\tau_{\mathrm{h}}$ and a selected jet.
- $m_{\ell\tau}$: mass of a lepton and the leading $\tau_{\mathrm{h}}$.
- $m_{\tau\tau}^{\mathrm{vis}}$: mass of the hadronic tau pair.

Table 5.13: Variables used in the event-level BDT classifiers in each event category to discriminant the $t\bar{t}H$ signal against $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}$ + jets background processes.

| Category | $1\ell + 2\tau_h$ | $2\ell$ss + $1\tau_h$ | $2\ell + 2\tau_h$ | $3\ell + 1\tau_h$ | $2\ell$ss | | $3\ell$ | |
|---|---|---|---|---|---|---|---|---|
| | | | | | $t\bar{t}$ | $t\bar{t}V$ | $t\bar{t}$ | $t\bar{t}V$ |
| cone-$p_T^{\ell_1}$ | ✓ | | ✓ | ✓ | | ✓ | | ✓ |
| cone-$p_T^{\ell_2}$ | | ✓ | | ✓ | | ✓ | | ✓ |
| max($\lvert\eta^{\ell}\rvert$) | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| $p_T^{\tau_1}$ | ✓ | ✓ | ✓ | ✓ | | | | |
| $p_T^{\tau_2}$ | ✓ | | ✓ | | | | | |
| $\lvert\cos\theta^{\tau_1}\rvert$ | ✓ | | ✓ | | | | | |
| $p_T^{miss}$ | ✓ | ✓ | | ✓ | | | | |
| $\Delta R(\ell_1, j)$ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $\Delta R(\ell_2, j)$ | | ✓ | | | ✓ | ✓ | ✓ | ✓ |
| $\Delta R(\ell_1, \ell_2)$ | | ✓ | | ✓ | | | | |
| $\Delta R(\tau_1, \tau_2)$ | ✓ | | ✓ | | | | | |
| $\Delta R(\tau_1, j)$ | ✓ | ✓ | | ✓ | | | | |
| $\Delta R(\tau_2, j)$ | ✓ | | | | | | | |
| $\Delta R(\ell_1, \tau_1)$ | ✓ | ✓ | | | | | | |
| $\Delta R(\ell_2, \tau_1)$ | | ✓ | | | | | | |
| $\Delta R_{ss}(\ell, \tau)$ | ✓ | | | | | | | |
| min($\Delta R(\tau, j)$) | | | ✓ | | | | | |
| $\langle\Delta R_{jj}\rangle$ | ✓ | ✓ | ✓ | | | | | |
| $\langle\Delta R_{\ell\tau}\rangle$ | | | ✓ | | | | | |
| $m_T^{\ell_1}$ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ |
| $m_T^{\ell_2}$ | | ✓ | | | | | | |
| $m_{\ell_1\tau}$ | | | | ✓ | | | | |
| $m_{\ell_2\tau}$ | | ✓ | | ✓ | | | | |
| $m_{\tau\tau}^{vis}$ | ✓ | | ✓ | | | | | |
| $m_{bb}$ | | ✓ | | ✓ | | | | |
| $q_{\ell_1\ell_2}$ | | | ✓ | | | | | |
| $N_j$ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| $N_b$ | ✓ | | ✓ | | | | | |
| $\mathcal{D}_{t_{had}}$ | ✓ | ✓ | | | ✓ | | | |
| Hadronic top $p_T$ | ✓ | ✓ | | | | | | |
| $\mathcal{D}_{Hj}$ | | | | | | ✓ | | |
| $\mathcal{D}_{MEM}$ | | | | | | | ✓ | ✓ |

- $m_{bb}$: mass of the two leading b-tagged jets. In the $2\ell$ss($3\ell$) + $1\tau_{\mathrm{h}}$ category, the medium (loose) working point is used. In case there are no such b-tagged jet pair, this variable is set to -1. Other event categories do not use this input variable.
- $m_{\mathrm{T}}^{\ell}$: transverse mass of a lepton. The lepton's transverse mass is defined as $m_{\mathrm{T}}^{\ell} = \sqrt{2p_{\mathrm{T}}^{\ell}p_{\mathrm{T}}^{\mathrm{miss}}(1-\cos\Delta\phi)}$, where $\Delta\phi$ is the difference in azimuthal angles between the lepton and the missing transverse momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$.
- $q_{\ell_1\ell_2}$: relative sign of the leading and trailing lepton charges.
- $\mathrm{N}_j$: number of selected jets.
- $\mathrm{N}_b$: number of loose b-tagged jets.
- $\mathcal{D}_{\mathrm{t_{had}}}$: a BDT-based hadronic top tagger. This BDT classifier computes the likelihood of a jet triplet to be compatible with hadronic top decay products. All possible combinations of jet triplets in an event are evaluated. For each jet triplet, one jet is labeled as the b jet candidate, and the other two are labeled as jets produced from hadronic W boson decay. The classification uses information including kinematics, b tagging discriminant of the b jet candidate, and quark-gluon likelihood discriminant of the W jet candidates. The jet triple with the highest BDT output is labeled as the hadronic top ($\mathrm{t_{had}}$) candidate. The discriminant value of this jet triplet is used as an input variable to the event-level BDT. The transverse momentum of the reconstructed hadronic top candidate is used in the event-level BDT as well.
- $\mathcal{D}_{\mathrm{Hj}}$: a BDT-based classifier to quantify the likelihood of a jet originating from H $\rightarrow$ WW decay (Higgs jet or Hj), in which one W boson decays hadronically. This classifier exploits jet kinematic and identification properties including b tagging and quark-gluon jet likelihood discriminants. It

also uses information regarding angular separations between the jet and leptons in the event. Each jet in an event is evaluated excluding those in the jet triplet tagged as the hadronic top candidate. The highest BDT output is used as an input variable to distinguish $t\bar{t}H$ signal from $t\bar{t}V$ background in the $2\ell$ss category.

- $\mathcal{D}_{\text{MEM}}$: discriminants based on the matrix element method [84–87]. The matrix element method combines theoretical inputs on signal and background cross sections with experimental effects such as detector resolutions. More detailed discussions on the implementation of MEM discriminants specifically in the context of $t\bar{t}H$ multi-lepton analysis can be found in Appendix A of Ref. [26].

Extra variables that are sensitive to tau helicity states may be added to further improve the signal extraction. Discussions on these variables can be found in Appendix B.

### 5.5.2 Statistical analysis

The statistical treatment follows procedures documented in Ref. [88]. The predicted distributions of the signal process, scaled by $\mu$, together with background processes are fitted to distributions obtained from data using a binned maximum likelihood (ML) fit.

**Maximum likelihood fit**

The likelihood function $\mathcal{L}$ is defined as

$$\mathcal{L}(\text{data} \mid \mu, \boldsymbol{\theta}) = \text{Poisson}\,(\text{data} \mid \mu, \boldsymbol{\theta}) \cdot \rho(\boldsymbol{\theta}|\tilde{\boldsymbol{\theta}}). \tag{5.13}$$

In the above equation, "data" can be either recorded data from experiments or toy data generated from statistical models. The parameter $\mu$ is the signal strength modifier and is the parameter of interest (POI). Systematic uncertainties are accounted for via nuisance parameters, the full collection of which is represented by $\boldsymbol{\theta}$ in Eq. 5.13.

For a binned likelihood, the term Poisson (data $\mid \mu, \boldsymbol{\theta}$) is a product of Poisson probabilities

$$\text{Poisson}\,(\text{data} \mid \mu, \boldsymbol{\theta}) = \prod_i \frac{(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i}}{n_i!} e^{-\mu s_i(\boldsymbol{\theta}) - b_i(\boldsymbol{\theta})}. \tag{5.14}$$

The number of events in bin $i$ of the discriminating observable distribution is denoted as $n_i$ in the above equation, and the product includes all bins in all event categories. The predicted signal and background yields in each bin $i$ are represented by $\mu s_i(\boldsymbol{\theta})$ and $b_i(\boldsymbol{\theta})$, respectively.

The function $\rho(\boldsymbol{\theta}|\tilde{\boldsymbol{\theta}})$ in Eq. 5.13 is the probability density function (*pdf*) of nuisance parameters. It parameterized all sources of systematic uncertainties. The parameter $\tilde{\boldsymbol{\theta}}$ represents the collection of the best estimated value of these nuisance parameters. These uncertainties are assumed to be either fully correlated or uncorrelated,[1] so that $\rho(\boldsymbol{\theta}|\tilde{\boldsymbol{\theta}})$ can be factorized into *pdf*s of individual nuisance parameters:

$$\rho(\boldsymbol{\theta}|\tilde{\boldsymbol{\theta}}) = \prod_k \rho(\theta_k|\tilde{\theta}_k). \tag{5.15}$$

[1] Partially correlated uncertainties are either factorized into fully correlated and uncorrelated sub-components, or are treated as fully correlated or uncorrelated, whichever is more appropriate or conservative.

The best estimated value $\tilde{\theta}_k$ of a nuisance parameter $\theta_k$ for a certain type of uncertainty $k$ is usually obtained from some auxiliary measurements or calculations.

The choice of a nuisance parameter *pdf* depends on the type of the uncertainty. For an uncertainty that only affects the overall normalization of a process, a log-normal distribution is used as the nuisance parameter's *pdf*, unless the uncertainty is of statistical origin, in which case a Gamma distribution is used instead. One example of the latter scenario is when event yields of a process are estimated by extrapolating from a sideband with limited amount of data. There could also be nuisance parameters that are unconstrained by any prior assumptions. In this case, a log-uniform distribution is used. In addition, some uncertainties could distort the shape of the discriminating observable distribution. These uncertainties, represented by Gaussian *pdf*s, are incorporated and evaluated in the likelihood function using the vertical morphing technique. A detailed description of the vertical morphing technique can be found in Section 4 of Ref. [89].

The best fit signal strength modifier is determined by maximize the likelihood function $\mathcal{L}(\text{data} \mid \mu, \boldsymbol{\theta})$. A test statistic $q_\mu$ can be defined based on the log likelihood ratio:

$$q_\mu = -2 \ln \frac{\mathcal{L}(\text{data} \mid \mu, \hat{\boldsymbol{\theta}}_\mu)}{\mathcal{L}(\text{data} \mid \hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{5.16}$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the maximum likelihood estimators of $\mu$ and $\boldsymbol{\theta}$, and $\hat{\boldsymbol{\theta}}_\mu$ maximize the likelihood function for a given $\mu$. By definition, $q_\mu$ is at its minimum value 0 when $\mu$ is the best fit value. Given that the test statistic $q_m u$ is expected to approximately following a $\chi^2$ distribution, the one sigma uncertainty of the POI $\mu$ is obtained by varying around its best fit value until the test statistic $q_\mu$

increases to 1.

**Binning strategy**

Distributions of discriminating observables in all event categories are binned for signal extractions. The binning strategy is individually optimized in each category, aiming to improve the analysis sensitivity while ensure that the relative statistical uncertainty of total background yields in each bin does not exceed 30%.

As mentioned in Section 5.5, two event-level BDTs are used for the signal extraction in the $2\ell$ss or $3\ell$ category. The 2D distribution of BDT outputs in either category is converted into 1D binned distribution before combining with other categories in the maximum likelihood fit. A binning map is used in each category to determine the bin number in the 1D distribution given the two BDT outputs of an event. The binning map is produced using simulated training events. A 2D signal-over-background likelihood ratio distribution is obtained by dividing the normalized 2D BDT distribution of $t\bar{t}H$ events by that of major backgrounds. The background distribution on the denominator contains events from simulated $t\bar{t}V$ and $t\bar{t}$+jets processes, weighted by their relative event counts after event selections in the corresponding event category. A 1D likelihood ratio distribution for backgrounds can then be obtained by looping over all background events and looking up their corresponding likelihood ratios. The binning is determined based on quantiles of the background 1D likelihood ratio distribution, and the binning map is generated by associating these quantiles to boundaries in the 2D phase space of the BDT outputs. The number of bins is set to eleven and six for $2\ell$ss and $3\ell$ categories, respectively.

In the $2\ell$ss+$1\tau_h$ category, the discriminating observable distribution is binned into eleven bins based on quantiles of the non-prompt background distribution. The binning in the $3\ell+1\tau_h$ category is based on quantiles of the total background distribution, and the number of bins is set to six in this category. Equal sized bins are used in event categories with two $\tau_h$'s. The number of bins are set to seven for $1\ell + 2\tau_h$ and four for the $2\ell + 2\tau_h$ category.

**Subcategory**

Because there are enough number of events selected in $2\ell$ss and $3\ell$ categories, the sensitivity of the analysis can be improved by further splitting the events into sub-categories based on lepton flavors (e or $\mu$), charges (+ or $-$), and b tagging requirements ("b-tight" (bt) if there are at least two medium b-tagged jets, "b-loose" (bl) otherwise). Doing so takes advantage of different signal-to-background ratios among these sub-categories. In general, identification efficiencies are lower and misidentification rates are higher for electrons than muons. Several background processes produced in proton-proton collisions, such as $t\bar{t}W$, WZ, W+jets, and signal top productions, more often contain positively charged leptons while this is not true for $t\bar{t}H$ signal events. In addition, there are generally fewer reconstructed b-tagged jets in non-prompt backgrounds due to lepton misidentification compared to $t\bar{t}H$ signal events.

**Treatment of $t\bar{t}W$ and $t\bar{t}Z$ backgrounds**

As discussed in Section 5.4.4, $t\bar{t}W$ and $t\bar{t}Z$ background processes are predicted based on MC simulated samples. The overall normalization of these processes

could be scaled to match the total amount of data used in the analysis based on their cross sections predicted by the standard model. However, both $t\bar{t}W$ and $t\bar{t}Z$ cross sections measured by CMS at 13 TeV exceed the standard model expectation by about one sigma [90]. In order to avoid potential bias to the $t\bar{t}H$ signal rate, the overall normalization of $t\bar{t}W$ and $t\bar{t}Z$ processes are allowed to scale freely by their strength modifiers $\mu_{t\bar{t}W}$ and $\mu_{t\bar{t}Z}$ in the fit. The $t\bar{t}W$ and $t\bar{t}Z$ enriched control regions are included in the signal extraction to better constrain these two parameters, and $\mu_{t\bar{t}W}$ and $\mu_{t\bar{t}Z}$ are determined simultaneously with the $t\bar{t}H$ signal strength modifier $\mu_{t\bar{t}H}$ from the maximum likelihood fit.

The $t\bar{t}WW$ background has a small contribution and is assumed to scale the same way as the $t\bar{t}W$ background.

## 5.6 Systematic uncertainties

Several theoretical and experimental uncertainties can lead to imperfect modelings of signal and background processes, varying either the overall normalization or the shape of the discriminating variable distribution. These systematic uncertainties are incorporated in the signal extraction as nuisance parameters in the likelihood function discussed in Section 5.5.2.

### Signal process

The standard model expected $t\bar{t}H$ cross section is relevant for the result of this analysis, since the $t\bar{t}H$ cross section is measured in terms of the signal strength modifier. The theoretical uncertainty associated with the expected $t\bar{t}H$ cross section, calculated at NLO accuracy, is $^{+6.8\%}_{-10.0\%}$. This uncertainty accounts for the $^{+5.8\%}_{-9.3\%}$ uncertainty due to missing higher order calculation in the perturbative expansion as well as 3.6% from uncertainties in parton distribution functions (PDFs) and $\alpha_s$ [4]. Uncertainties in the Higgs boson branching ratios to a pair of vector bosons (W or Z) and a pair of $\tau$ leptons amount to 1.5% and 1.7% [4], respectively. Uncertainties in renormalization and factorization scales are accounted for by varying these scales independently by a factor of 2 and 1/2 with respect to their nominal values [4].

### Background processes

No theoretical uncertainty is necessary for the $t\bar{t}Z$ and $t\bar{t}W(W)$ backgrounds, since their contributions are determined from the fit. Effects on the shape of the

discriminating observable for $t\bar{t}Z$ and $t\bar{t}W(W)$ backgrounds due to uncertainties of theoretical origins are negligible.

A 100% uncertainty is assigned to the WZ cross section in all event categories except $3\ell$. This conservative value is chosen because it is concerned that the accurate theoretical calculation of WZ cross section, supplemented by measurements from both ATLAS and CMS at the LHC [91, 92], does not necessarily apply to the $t\bar{t}H$ signal region, in which the presence of multiple light and heavy flavored jets is required. As for the $3\ell$ category, in which the WZ process has a larger contribution, its normalization is studied more carefully in the WZ control region. The WZ contribution in the $3\ell$ signal region is constrained based on the control region extrapolation. An overall uncertainty of 50% is assigned to it in this category.

Uncertainties in the non-prompt background arise from uncertainties in the measured fake factors as well as imperfect extrapolations using the Fake Factor method. The latter is evaluated based on the closure test as described in Section 5.4.1. Uncertainties in the fake factor measurement include the statistical uncertainty of the sample used in the measurement as well as estimations of prompt contamination in the DR. Both the normalization and the shape of the misidentified background are affected by these uncertainties. The normalization uncertainty is 50% in $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ categories, and is about 30% in other event categories with zero or one $\tau_h$. Effects on the shape of the discriminating observable distribution are evaluated by varying the fake factors independently. In $2\ell + 1\tau_h$ and $3\ell + 1\tau_h$ categories, the data-driven misidentified background estimation using the FF method only accounts for background events due to misidentification of leptons, and the background due to misiden-

tification of $\tau_h$'s is estimated in simulated events based on generator level object matching. As a result, an additional 30% uncertainty on $\tau_h$ fake factors in simulation is assigned to the $t\bar{t}H$ and irreducible background MC events that contain a misidentified $\tau_h$.

A 50% uncertainty is assigned to other minor irreducible background processes estimated using MC simulation. Uncertainties of photon conversion background and the background due to electron charge misidentification in $2\ell$ss and $2\ell$ss + $1\tau_h$ categories both amount to 30%.

**Trigger efficiency**

Uncertainties associated with trigger efficiency scale factors for event categories using lepton triggers are listed in Table 5.11. The trigger scale factor uncertainty varies between $1-3$% depending on the number of leptons, lepton flavors, and leading lepton $p_T$ in an event. In the $1\ell + 2\tau_h$ category, efficiency sale factors of single-lepton triggers and lepton-$\tau_h$ cross triggers are measured to have 3% uncertainties.

**Lepton and $\tau_h$ selection efficiency**

The uncertainty in the efficiency of a muon passing the tight selection is about 3%. For an electron, the selection efficiency has an uncertainty of 3% if its $p_T$ is greater than 25 GeV, and 5% uncertainty otherwise. An uncertainty of 5% is assigned to $\tau_h$ reconstruction and identification efficiencies [63].

**Jet and $\tau_h$ energy scale**

The uncertainty in the jet energy scale is a few percent depending on $p_T$ and $\eta$ of the jet. The $\tau_h$ energy scale is measured with an uncertainty of 1.2% [63]. Both jet and $\tau_h$ energy scale uncertainties change the shape of the discriminating observable distribution. Their effects are evaluated by varying energies of all jets or $\tau_h$'s within their corresponding energy scale uncertainties and reprocessing all events. The change of energy scale is also propagated to the calculation of missing transverse momentum.

**b tagging efficiency and mistag rate**

Uncertainties in b tagging efficiency and mistag rate are split into individual sources and are parameterized as a function of jet $p_T$ and $\eta$ [60]. Their impacts on the analysis result are evaluated by separately varying the corresponding data-to-MC scale factors.

**Integrated luminosity**

The uncertainty on the integrated luminosity for data recorded in 2017 is 2.3% [93].

Major systematic uncertainties, as well as their impact on the fitted $t\bar{t}H$ signal strength modifier, are summarized in Table 5.14. The impact on signal strength modifier ($\Delta\mu/\mu$) is the relative variation of the signal rate due to shifting the systematic source by its corresponding uncertainty. Impacts on the signal strength modifier for the analysis based on 2017 data set as well as for 2016 and 2017

Table 5.14: Summary of the main sources of systematic uncertainty and their impact on the measured signal rate.

| Source | Uncertainty [%] | $\Delta\mu/\mu$ [%] (2017) | $\Delta\mu/\mu$ [%] (Comb.) | Correlated |
|---|---|---|---|---|
| Theoretical sources | $\approx 8$ | 8 | 9 | Yes |
| e, $\mu$ selection efficiency | 3–5 | 4 | 3 | Yes |
| $\tau_h$ selection efficiency | 5 | 3 | 5 | Yes |
| $\tau_h$ energy calibration | 1.2 | 1 | 2 | Yes |
| b tagging efficiency | 2–15 | 10 | 5 | Partially |
| Jet energy calibration | 2–15 | 3 | 3 | Yes |
| Fake background yield | $\approx$ 30–50 | 17 | 9 | Partially |

combined analysis are shown in the table.

For the purpose of combining the 2016 and 2017 data sets, most systematic uncertainties are considered to be correlated between the two years. A few exceptions include uncertainties in trigger efficiency, b tagging, integrated luminosity, and non-prompt background rate. For b tagging uncertainties, only non-statistical sources are correlated while others are treated as uncorrelated between the two years. Regarding uncertainties in non-prompt backgrounds, contributions from uncertainties in lepton fake factor measurements are treated as correlated, while uncertainties from closure tests are uncorrelated. The uncertainty in the $\tau_h$ fake factor measurement is dominated by statistical uncertainties due to the limited number of data events in the determination region. Therefore, it is considered to be uncorrelated between the two years. The correlation treatment of major sources of systematic uncertainties are indicated in the last column of Table 5.14.

## 5.7 Results

The $t\bar{t}H$ signal strength modifier $\mu$ is determined by a maximum likelihood fit using data recorded in 2017, corresponding to an integrated luminosity of 41.5 $fb^{-1}$. Distributions of the discriminating observable in all event categories (except $4\ell$) after the fit are shown in Fig. 5.7 and 5.9. Post-fit distributions of the discriminating observable in the $t\bar{t}W$ and $t\bar{t}Z$ control regions are also shown in Fig. 5.8. Event yields in all categories after the fit are summarized in Table 5.15. In each category, the total number of expected events, indicated as "SM expectation" in the table, is computed by summing up predicted yields of both signal and background processes. The expected total number of events in each category is consistent with the observed number of events from data within the uncertainties.

The observed (expected) $t\bar{t}H$ signal strength modifier extracted from the ML fit is $\mu = 0.75^{+0.46}_{-0.43}$ ($1.00^{+0.39}_{-0.35}$), combining all seven event categories. The observed $\mu$ agrees with the standard model expectation within its uncertainties. The observed (expected) significance of the excess over background-only hypothesis (i.e. no $t\bar{t}H$) is $1.7\sigma$ ($2.9\sigma$).

The $t\bar{t}H$ signal strength modifier can also be extracted in each event category individually by allowing different $\mu$ values in different categories. The best fit $\mu$ values in individual event categories, together with the result combining all categories, are shown in Fig. 5.10 and Table 5.16. For the fit in the $2\ell + 2\tau_h$ category, a lower boundary of zero on the $\mu$ range is enforced to avoid the divergence in the fit due to very low number of entries in some bins of the discriminating observable distribution. The observed signal strength modifiers in individual

categories are compatible with the SM expectation and with each other. As indicated in Fig. 5.10 and Table 5.16, the sensitivity of the analysis is mostly driven by $2\ell$ss, $3\ell$, and $2\ell$ss + $1\tau_h$ categories, while $2\ell + 2\tau_h$, $3\ell + 1\tau_h$, and $4\ell$ categories are statistically limited.

As discussed in Section 5.5.2, the overall normalization of $t\bar{t}W$ and $t\bar{t}Z$ backgrounds are parameterized by $\mu_{ttW}$ and $\mu_{ttZ}$, respectively. They are determined in the fit simultaneously with the $t\bar{t}H$ signal strength modifier. The observed (expected) values are $\mu_{t\bar{t}W} = 1.42^{+0.34}_{-0.33}$ $(1.00^{+0.27}_{-0.24})$ and $\mu_{t\bar{t}Z} = 1.69^{+0.39}_{-0.33}$ $(1.00^{+0.24}_{-0.21})$, both of which follows the same trend as reported in the measurement by CMS on $t\bar{t}W$ and $t\bar{t}Z$ cross sections [90].

The result on $t\bar{t}H$ cross section measurement based on the 2017 data set is also combined with that based on the 2016 data set, corresponding to an integrated luminosity of 35.9 $fb^{-1}$. The strategy of the analysis with the 2016 data set is similar to the analysis described in this thesis. More details on the analysis with 2016 data set can be found in Ref. [26]. Correlations of systematic uncertainties between data sets of the two years are discussed at the end of Section 5.6. The normalization of $t\bar{t}W$ and $t\bar{t}Z$ backgrounds are allowed to freely float and are determined from the fit. The combination of 2016 and 2017 results yields an observed (expected) $t\bar{t}H$ signal strength $\mu = 0.96^{+0.34}_{-0.31}$ $(1.00^{+0.30}_{-0.27})$. The observed (expected) significance of the $t\bar{t}H$ signal is $3.2\sigma$ $(4.0\sigma)$. The combined result agrees very well with the SM predicted $t\bar{t}H$ production cross section.

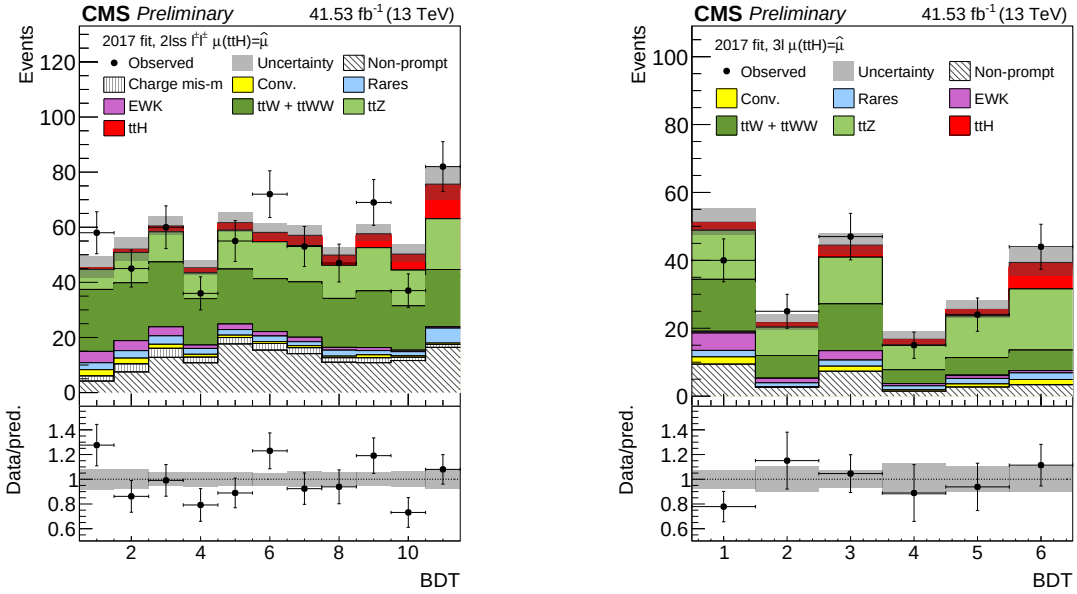


Figure 5.7: Distributions of discriminating observables after the fit in $2\ell$ss (left) and $3\ell$ (right) categories.

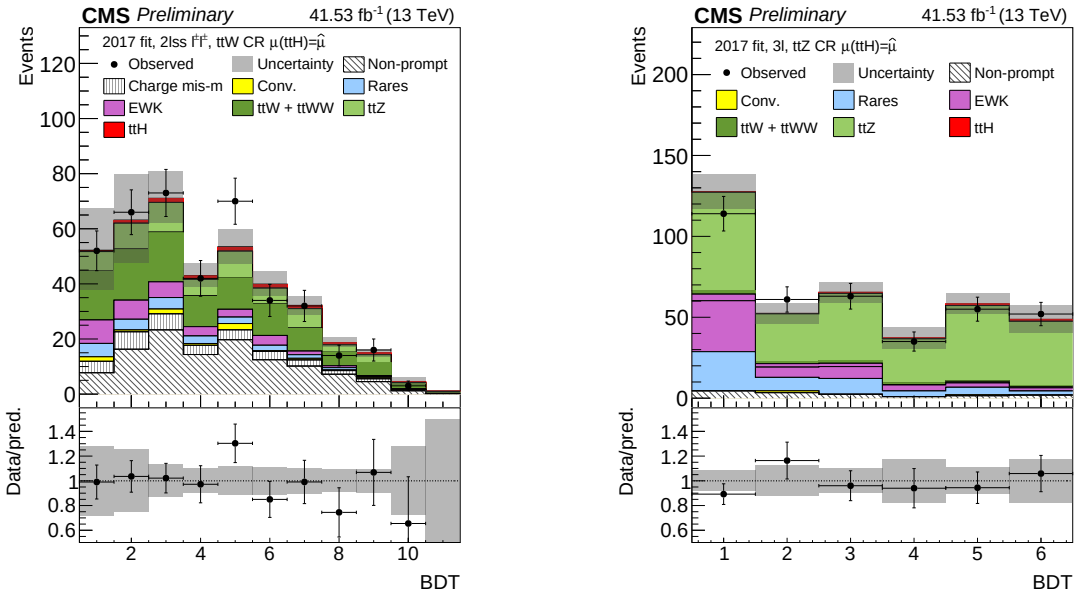


Figure 5.8: Distributions of discriminating observables after the fit in $t\bar{t}W$ (left) and $t\bar{t}Z$ (right) control regions.

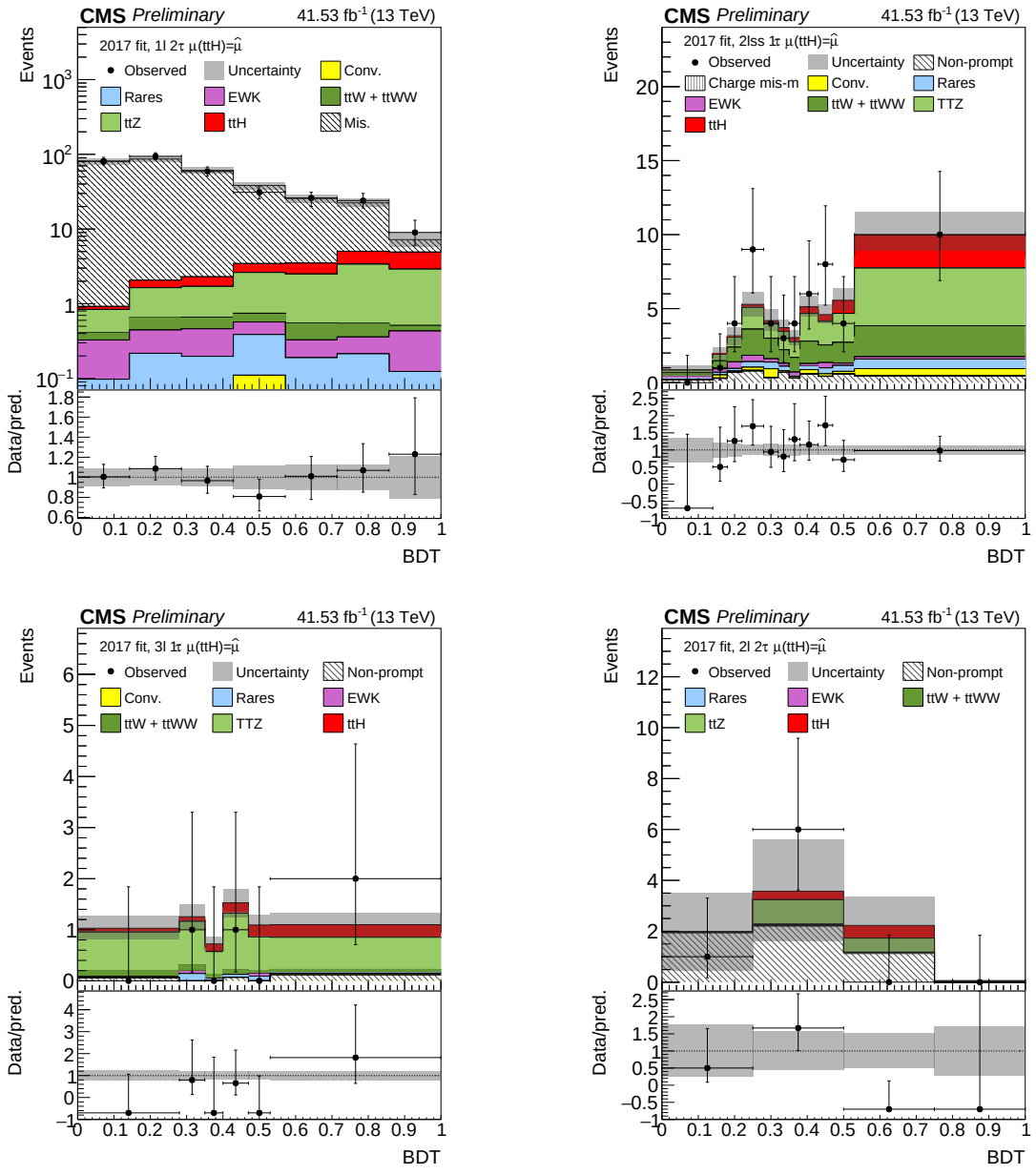


Figure 5.9: Post-fit distributions of discriminating observables in $1\ell + 2\tau_{\mathrm{h}}$ (top left), $2\ell\mathrm{ss} + 1\tau_{\mathrm{h}}$ (top right), $3\ell + 1\tau_{\mathrm{h}}$ (bottom left), and $2\ell + 2\tau_{\mathrm{h}}$ (bottom right) categories.

Table 5.15: Event yields in all categories after the maximum likelihood fit. The "SM expectation" in the table includes contributions from the $t\bar{t}H$ process.

| Category | $2\ell$ss | $3\ell$ | $4\ell$ | $1\ell + 2\tau_h$ | $2\ell + 2\tau_h$ | $3\ell + 1\tau_h$ | $2\ell$ss $+ 1\tau_h$ |
|---|---|---|---|---|---|---|---|
| $t\bar{t}H$ | 43.0 ± 7.1 | 18.8 ± 4.8 | 0.7 ± 0.3 | 6.6 ± 3.6 | 0.9 ± 0.5 | 1.0 ± 0.4 | 5.1 ± 2.1 |
| $t\bar{t}W + t\bar{t}WW$ | 218.5 ± 13.7 | 51.0 ± 5.3 | 0.13 ± 0.03 | 1.1 ± 0.3 | < 0.05 | 0.5 ± 0.1 | 13.1 ± 2.4 |
| tH | 2.4 ± 0.1 | 0.9 ± 0.1 | < 0.05 | 0.3 ± 0.1 | < 0.05 | < 0.05 | 0.5 ± 0.0 |
| WZ + ZZ | < 0.05 | 12.0 ± 1.7 | 0.15 ± 0.10 | 1.5 ± 0.8 | < 0.05 | 0.1 ± 0.0 | 2.8 ± 2.0 |
| $t\bar{t}Z/\gamma^*$ | 138.2 ± 7.6 | 74.1 ± 6.3 | 3.9 ± 0.6 | 11.6 ± 2.6 | 1.6 ± 0.5 | 4.5 ± 0.7 | 15.4 ± 2.4 |
| Misidentified | 132.1 ± 10.0 | 26.8 ± 4.0 | < 0.05 | 299.6 ± 19.1 | 5.3 ± 2.2 | 0.3 ± 0.3 | 5.3 ± 2.2 |
| Conversions | 11.6 ± 3.0 | 6.6 ± 1.3 | < 0.05 | 0.3 ± 0.1 | < 0.05 | < 0.05 | < 0.05 |
| Charge flip | 22.8 ± 2.3 | < 0.05 | < 0.05 | < 0.05 | < 0.05 | < 0.05 | < 0.05 |
| Other | 26.7 ± 3.9 | 9.7 ± 2.2 | < 0.05 | 1.2 ± 0.5 | 0.06 ± 0.04 | 0.3 ± 0.2 | 3.2 ± 1.1 |
| SM expectation | 595.3 ± 20.6 | 200.0 ± 10.8 | 5.0 ± 0.7 | 322.0 ± 19.6 | 7.9 ± 2.3 | 6.7 ± 0.9 | 45.3 ± 5.1 |
| Observed data | 614 | 195 | 6 | 324 | 7 | 4 | 53 |

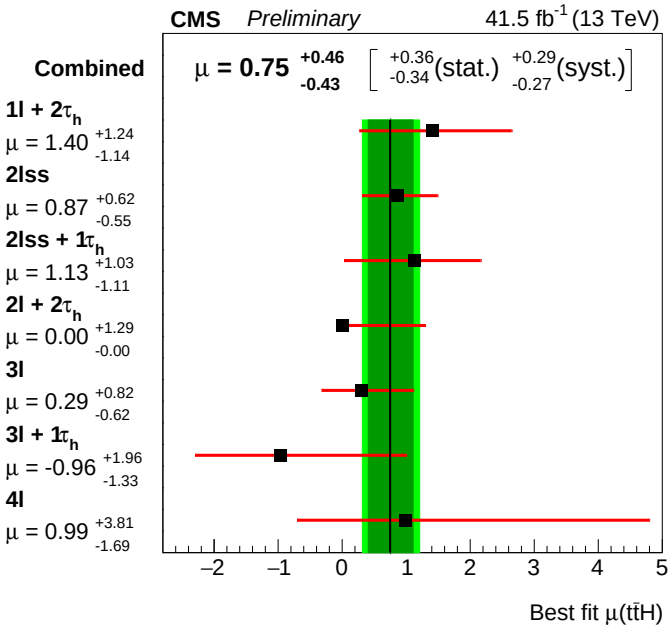


Figure 5.10: Measured $t\bar{t}H$ signal strength modifiers for individual event categories and for the combination of all event categories using 41.5 $fb^{-1}$ data recorded in 2017.

Table 5.16: Measure and expected $t\bar{t}H$ signal strength modifiers in each event category and with all seven categories combined using data recorded in 2017. The 2017 result is also combined with that obtained using data recorded in 2016. The result combining two years is shown in the last row of the table.

| | Signal Strength $\pm1\sigma$ | |
|---|---|---|
| Category | Measured | Expected |
| $1\ell + 2\tau_h$ | $1.40^{+1.24}_{-1.14}$ | $1.00^{+1.14}_{-0.93}$ |
| $2\ell ss$ | $0.87^{+0.62}_{-0.55}$ | $1.00^{+0.53}_{-0.49}$ |
| $2\ell ss + 1\tau_h$ | $1.13^{+1.03}_{-1.11}$ | $1.00^{+0.93}_{-0.80}$ |
| $2\ell + 2\tau_h$ | $0.00^{+1.29}_{-0.00}$ | $1.00^{+2.63}_{-1.56}$ |
| $3\ell$ | $0.29^{+0.82}_{-0.62}$ | $1.00^{+0.59}_{-0.52}$ |
| $3\ell + 1\tau_h$ | $-0.96^{+1.96}_{-1.33}$ | $1.00^{+1.91}_{-1.37}$ |
| $4\ell$ | $0.99^{+3.31}_{-1.69}$ | $1.00^{+2.41}_{-1.72}$ |
| Combined 2017 | $0.75^{+0.46}_{-0.43}$ | $1.00^{+0.39}_{-0.35}$ |
| Combined 2016+2017 | $0.96^{+0.34}_{-0.31}$ | $1.00^{+0.30}_{-0.27}$ |

CHAPTER 6

# CMS LEVEL-1 TRACK TRIGGER UPGRADE FOR THE HL-LHC

This chapter presents a proposed hardware-based track trigger system for future CMS detector at the High-Luminosity LHC (HL-LHC). The upgraded detector is expected to maintain or exceed the current performance under the very challenging condition at the HL-LHC. Section 6.1 and 6.2 introduces the planned HL-LHC and the CMS Phase-2 upgrade, motivating the necessity of adding tracking information to the Level-1 trigger. Section 6.3 focuses on one approach, named "Tracklet", for realizing Level-1 tracking. The algorithm, implementation, performance, and a hardware demonstrator of the Tracklet approach are described in details in this section. A summary is given in Section 6.4.

## 6.1 High-Luminosity LHC

The LHC is going to operate for another three years from 2021 to 2023 (Run 3), before undergoing major upgrades during the so-called Long Shutdown 3 (LS3) starting from 2024. The upgraded collider complex, High-Luminosity LHC, will collide proton bunches at a center-of-mass energy of 14 TeV every 25 ns. The instantaneous luminosity of the HL-LHC is expected to be up to $7.5 \times 10^{34}$ $\text{cm}^{-2}\text{s}^{-1}$ [94], which corresponds to an average number of pileup up to 200. The LHC and HL-LHC timeline, together with collision energies and luminosities, is summarized in Fig. 6.1. By the end of the HL-LHC operation, the integrated luminosity is expected to be 3000 $\text{fb}^{-1}$, which is ten times that expected over the lifetime of the LHC. The HL-LHC will greatly benefit detailed studies of Higgs boson properties and also expand the physics potential

for discovering new physics beyond the standard model. However, the harsh condition at the HL-LHC, in particular the high pileup, makes it very challenging to build a detector that can fully exploit the delivered physics potential. In order for CMS to keep providing high quality data in the era of HL-LHC, all CMS subsystems will be substantially upgraded during LS3, referred to as the CMS Phase-2 Upgrade [95].


Figure 6.1: LHC/HL-LHC operation and upgrade timeline [96]. The center-of-mass energy and instantaneous luminosity during each period are indicated in the figure. The nominal instantaneous luminosity of the LHC is $1.0 \times 10^{34}$ cm$^{-2}$s$^{-1}$. The total amount of data in terms of integrated luminosities at the end of each operation period are also shown.

The HL-LHC pileup condition is especially challenging for the CMS trigger system to select data that are of interest for further studies. As mentioned in Section 3.2.7, CMS adopts a two-level trigger system: the hardware-based level-1 (L1) trigger currently using only information from calorimeters and the muon system, and the software-based high-level trigger (HLT). Under the HL-LHC pileup condition, the current L1 trigger would not be able to handle the high data rate from the front end, and its output trigger rate would be unsustainably high. Attempts to lower the trigger rate by significantly increasing trigger thresholds not only would limit the physics potential by reducing the trigger ef-

ficiency for the physics of interest, but also would be insufficient. The planned Phase-2 L1 trigger will have a higher L1 trigger maximum rate of 750 kHz (currently 100 kHz) and an increased latency budget of 12.5 $\mu$s (currently 3.2 $\mu$s). In addition, the charged particle tracking information will be integrated into the L1 trigger. Tracking at L1 trigger provides extra handles for managing the trigger rate in a high pileup environment. It also helps to improve $p_{\mathrm{T}}$ resolutions and identifications of various L1 objects, as well as providing additional vertex and track isolation information for triggering on hadronic activities.

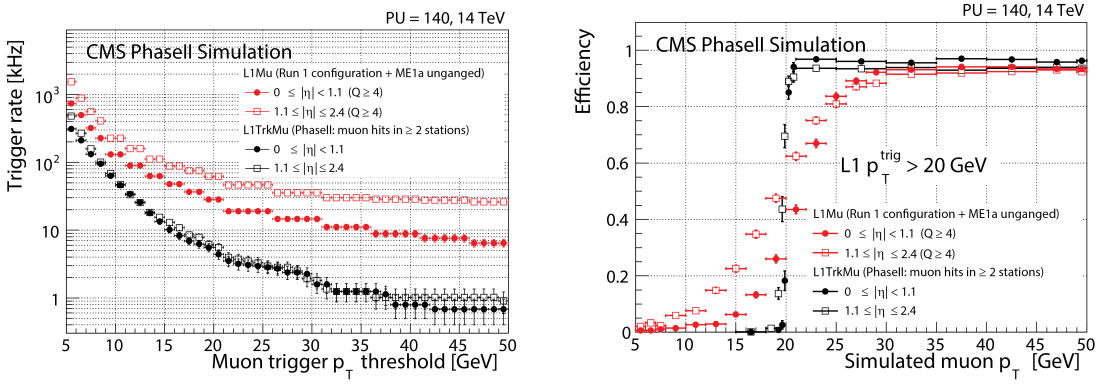


Figure 6.2: Left: L1 single muon trigger rate as a function of trigger $p_{\mathrm{T}}$ thresholds; Right: the efficiency of the L1 single muon trigger with a threshold $p_{\mathrm{T}} > 20$ GeV as a function of simulated muon $p_{\mathrm{T}}$ [95]. The L1 standalone muon trigger (red) and the trigger combining L1 muon and L1 track information (black) are shown in the pseudo-rapidity range of $|\eta| < 1.1$ (solid circles) and $1.1 \leq |\eta| \leq 2.4$ (hollow circles).

An example of the L1 single muon trigger performance with a pileup of 140 is shown in Fig. 6.2. The plot on the left shows the trigger rate as a function of trigger $p_{\mathrm{T}}$ thresholds, and the right plot shows the trigger efficiency of a L1 muon trigger with a $p_{\mathrm{T}}$ threshold of 20 GeV as a function of simulated muon $p_{\mathrm{T}}$. The data points in red correspond to the standalone L1 muon trigger with Run 1 configuration, denoted as L1Mu, and the black data points represent the same trigger but with additional L1 tracking information added to it, denoted as L1TrkMu. As shown in the left plot, increasing the $p_{\mathrm{T}}$ threshold of the stan-

dalone L1 muon trigger eventually becomes ineffective to reduce the trigger rate, because it is more difficult for the trigger to determine the muon momentum as it becomes larger. With additional tracking information available at L1, a better $p_{\mathrm{T}}$ measurement leads to a more effective reduction of the trigger rate. For a trigger with a chosen threshold, an improved $p_{\mathrm{T}}$ resolution provided by the L1 tracking means a quicker turn-on of the trigger. As shown in the right plot of Fig. 6.2, almost all events containing muons with $p_{\mathrm{T}}$ less than the 20 GeV threshold are rejected by the L1TrkMu trigger, while events with muon $p_{\mathrm{T}}$ greater than the threshold are selected by this trigger with very high efficiencies.

## 6.2 CMS Phase-2 tracker

The layout and module design of the Phase-2 outer silicon tracker is largely driven by the necessity of providing tracking information to the L1 trigger. To make tracking at L1 feasible, the data volume sent out at the 40 MHz collision rate has to be limited. This is achieved with a novel module design, referred to as "$p_{\mathrm{T}}$ modules" [97], which can provide local data reductions at the front end. The $p_{\mathrm{T}}$ module concept is illustrated in the top plot of Fig. 6.3. A $p_{\mathrm{T}}$ module consists of two single-sided closely-spaced silicon sensors, which are read out by a common set of front-end electronics. Discrimination on $p_{\mathrm{T}}$ are achieved in the $p_{\mathrm{T}}$ module by correlating hits in the two sensors based on the bend of the hit pair in the CMS magnetic field. Only those pairs, referred to as "stubs", that are compatible with particle tracks above a configurable $p_{\mathrm{T}}$ threshold are sent out. By doing so, a factor of 10 to 100 local data reduction can be achieved in the front-end electronics. There are two types of $p_{\mathrm{T}}$ modules used in the CMS Phase-2 outer tracker: strip-strip (2S) modules and pixel-strip (PS) modules.

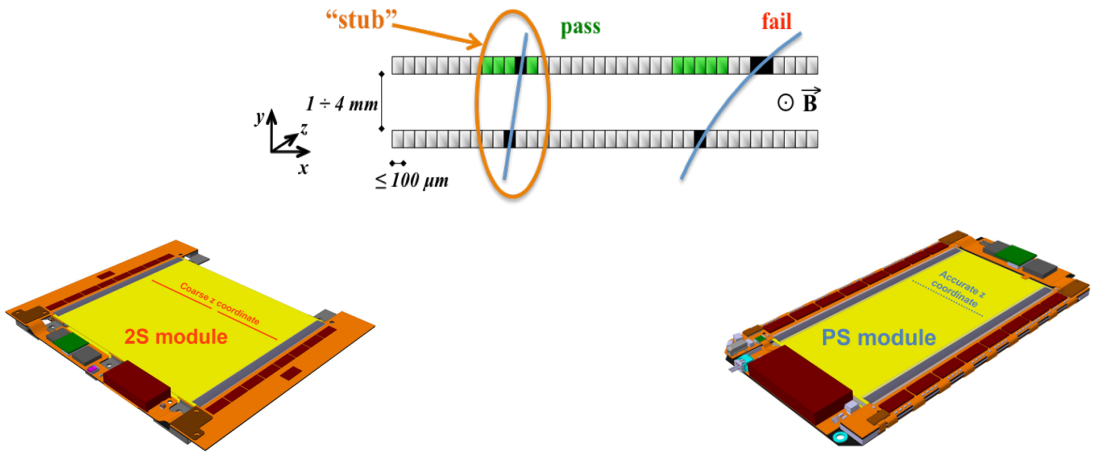


Figure 6.3: Top: the concept of $p_{\mathrm{T}}$ modules. Hits in the two closely spaced sensors are correlated to provide on-detector $p_{\mathrm{T}}$ discrimination. Bottom: strip-strip modules (left) and pixel-strip modules (right). All plots are taken from Ref. [98].

The 2S and PS modules are shown in the bottom two plots of Fig. 6.3. The strip direction is parallel to the $z$ axis in the barrel and approximately radial in the endcaps. The 2S module has two $10 \times 10$ cm$^2$ strip sensors. Each strip sensor contains two rows of 5 cm long strips, and each row contains 1016 strips. The strip pitch is 90 $\mu$m for 2S modules. The PS module has a strip sensor on the top and a macro-pixel sensor on the bottom. The PS strip sensor is half the size of the 2S strip sensor, with two rows of 2.35 cm long strips. Each row contains 960 strips with a pitch of 100 $\mu$m. The macro-pixel sensor can provide more precise $z$ coordinate measurements. It contains $960 \times 32$ macro-pixels with a length of 1467 $\mu$m in $z$ and a pitch of 100 $\mu$m.

A sketch of one quarter of the CMS Phase-2 tracker layout in $r$-$z$ view is shown in Fig. 6.4. For the outer tracker, there are six barrel layers and five endcaps on each side. The red modules in the tracker layout are 2S modules. They are implemented in the outer three barrel layers and in the endcap region with radii larger than about 700 mm. The PS modules are represented by blue mod-

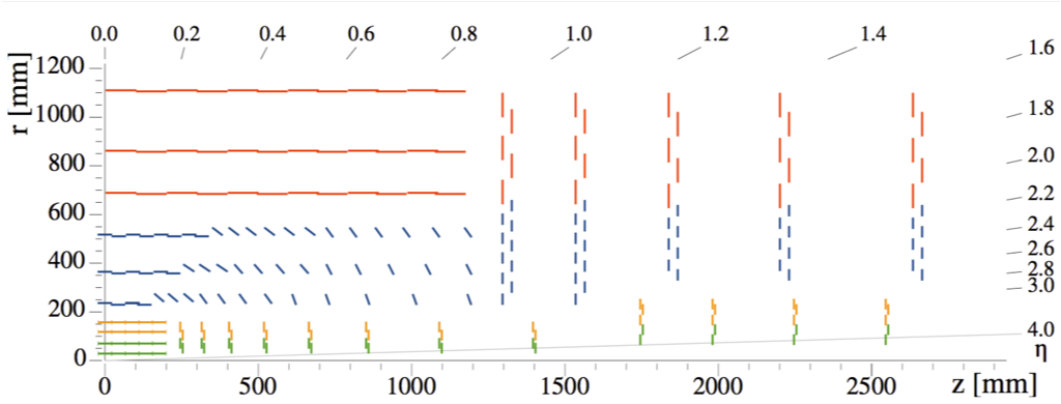


Figure 6.4: The layout of the proposed CMS Phase-2 tracker in the $r$-$z$ view [98]. One quarter of the tracker is shown in the plot. The blue and red modules are pixel-strip and strip-strip modules for the silicon strip tracker, respectively. The inner pixel tracker modules are indicated in orange and green, and they are not used in the L1 track finding.

ules and are implemented in the first three barrel layers of the outer tracker and in the endcap with smaller radii. The orange and green modules belong to the inner pixel tracker and are not used in the L1 track finding. As shown in Fig. 6.4, the PS modules at larger $z$ in the first three barrel layers are tilted towards the interaction point (IP). This so-called "titled geometry" is chosen to improve the front-end stub efficiency over the "flat geometry", in which all modules are parallel to the $z$ axis. The $p_{\mathrm{T}}$ module concept relies on the correlated readout of the top and bottom sensors at the same time. In practice, the two halves of each module are read out independently one the two ends. While the halves of the top and bottom sensors on the same side can communicate, they do not share information with the other side. In case that a charged particle crosses the module near its center with a large incident angle, hitting one half of the bottom sensor and the other half of the top sensor, no stub would be read out. This inefficiency is particularly prominent in the barrel layers at a large $\eta$. For this reason, some PS modules in the first three barrel layers are progressively tilted, so that they are approximately perpendicular to incident particles.

## 6.3 Tracklet approach for L1 tracking

The goal of the L1 track finding is to perform pattern recognition to identify those stubs that belong to the same charged particle trajectory with $p_{\mathrm{T}}$ higher than a configurable threshold (e.g. 2 GeV) and to compute the helix parameters of the trajectory. The L1 track finding system needs to be able to handle the high data throughput from the detector at the 40 MHz proton collision rate. Each proton bunch crossing at the HL-LHC would create about $15,000$ stubs, about 10% of which are associated with tracks of interest. In addition, the L1 track finder needs to reconstruct tracks within 4 $\mu$s in order to be used in the L1 trigger decision.

There are a few approaches considered by the CMS collaboration to implement the L1 track finder. This section focuses on one approach, referred to as the "Tracklet" [99], based on the field-programmable gate array (FPGA) technology. Other proposals include an FPGA-based Hough transform approach, referred to as "TMTT" [100], and an associative memory based approach using custom ASICs [101]. The commercially available FPGA is an ideal hardware platform for implementing fast tracking algorithms because of its ever-increasing capability and reconfiguration flexibility. The available computing and storage resources on FPGAs, in particular the digital signal processors (DSP) and random-access memories (RAM), are well suited for highly parallelized architectures that are necessary for L1 track finding.

### 6.3.1 Algorithm

The Tracklet approach adopts a "road search" algorithm. The Tracklet algorithm is briefly illustrated in Fig. 6.5. The algorithm is highly parallelized both in space, via partitioning the detector into smaller regions, and in time, via the time-multiplexed architecture. The tracker is partitioned into so-called "sectors" based on azimuth angles. Each sector is processed in parallel by one dedicated hardware unit containing one FPGA. For a time-multiplexing (TMUX) factor of $n$ (typically 6 - 18), $n$ independent duplicates of the system are made. Each unit processes a new proton bunch collision, or an "event", every $n \times 25$ ns.

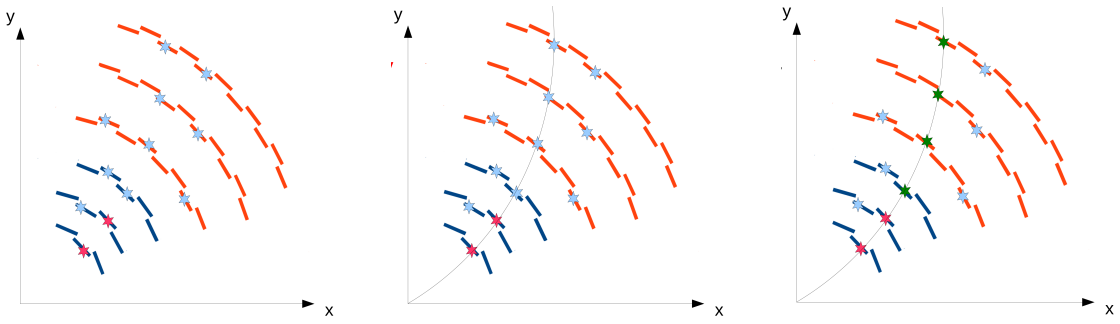


Figure 6.5: An overview of the Tracklet algorithm. Left: the seeding step, in which a tracklet is formed from a pair of stubs (red) in adjacent layers/disks. Middle: the projection step, in which projected positions to other layers/disks are estimated based on the seed. Right: the matching and fitting step, in which stubs in other layers/disks are compared to the projected trajectory. The matched stubs (green) and the seeding stubs (red) are grouped together to compute the track parameters by a $\chi^2$ fit.

#### Data organization

As mentioned briefly above, the detector is split equally into $\phi$ sectors. The number of $\phi$ sectors is chosen based on several factors including cabling configurations of the upstream system, the input bandwidth needed per sector, the TMUX factor, and the cost of FPGAs. In the current system, the number of sec-

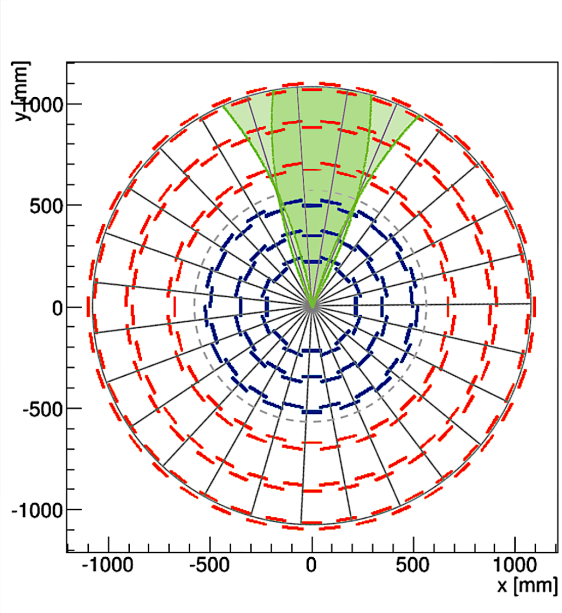


Figure 6.6: The definition of processing sectors in $x$-$y$ view of the tracker. The green highlighted area is one of the nine $\phi$ sectors. Dotted circle in gray indicated the critical radius $r_c$.

tors is nine, and the sector definition, referred to as "hourglass" configuration, is illustrated in Fig. 6.6. One of the nine sectors ("nonants") is highlighted in green in Fig. 6.6. The boundary of a sector are based on both positively and negatively charged particle tracks with $p_{\mathrm{T}} = 2$ GeV. The positively and negatively charged particle tracks cross at a configurable radius, referred to as the critical radius and denoted as $r_c$. Each $\phi$ sector covers the entire $z$ region of the tracker. Such sector definition ensures any $p_{\mathrm{T}} > 2$ GeV track would be either entirely inside the sector or not in the sector at all. In contrast to the previous pizza-shaped sector definition with straight line boundaries, indicated in gray in Fig. 6.6, each hourglass sector can be processed independently, and no inter-sector data transmission is needed. With the hourglass configuration, however, overlaps between the neighboring sectors are inevitable. Stubs in these overlap regions, indicated in light green in Fig. 6.6, have to be duplicated and sent to both sector processors. The critical radius $r_c$ is chosen to optimize for the

input cabling configuration and to reduce the number of stubs that need to be duplicated.

### Pattern recognition

The pattern recognition starts by forming seeds, or "tracklets", between two adjacent layers/disks. Seeds are formed in parallel in several layer/disk pairs including barrel layer 1 and 2 (L1L2), 3 and 4 (L3L4), 5 and 6 (L5L6), disk 1 and 2 (D1D2), 3 and 4 (D3D4), layer 1 and disk 1 (L1D1), layer 2 and disk 1 (L2D1). The tracking efficiency for each seeding combination as a function of $\eta$ is shown in Fig. 6.7. As can be seen in this plot, seeding in multiple layer/disk combinations provides sufficient coverage as well as redundancy in the tracker $\eta$ range.

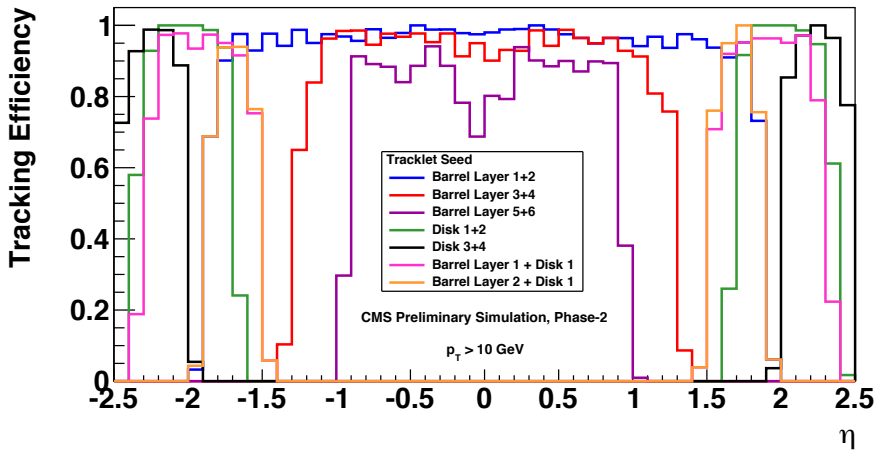


Figure 6.7: Tracking efficiency in single muon events as a function of $\eta$ for different seeding pairs.

To further increase the level of parallelization, each barrel layer or endcap disk in a sector is further divided based on $\phi$ into 16 or 32 smaller regions, referred to as virtual modules (VM). Each VM covers the full $z$ range. Track seeds are formed from stubs in one VM of the inner seeding layer/disk and stubs in

one VM of the outer seeding layer/disk. Each combination of inner and outer VM pairs is processed in parallel. With e.g. 32 VMs in layer 1 and 32 VMs in layer 2, in principle there would be $32^2 = 1024$ combinations to consider. However, only 154 VM pairs among them can possibly form seeds that are compatible with charged particles of $p_T > 2$ GeV produced near the interaction point. Therefore, dedicated processing units are assigned only to these 154 VM pairs. By doing so, a large reduction of combinatorics is achieved. The concept of virtual modules and how VMs help to reduce seeding combinations are illustrated in Fig. 6.8.

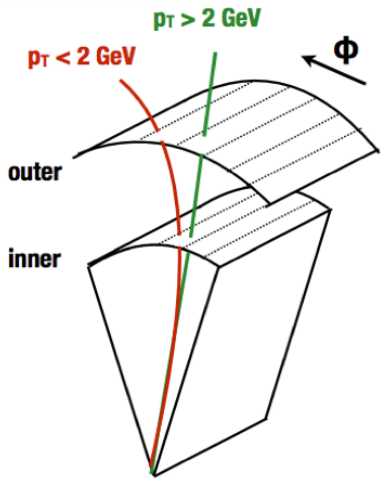


Figure 6.8: An illustration of virtual modules in a pair of seeding layers. A track with its $p_T$ below the threshold (red) would bend more than a track with a larger $p_T$ (green) in the CMS magnetic field. Therefore only some combinations of inner and outer VMs are worth spending computing resources to process.

With a pair of seeding stubs in the inner and outer VMs, a set of parameters for the helical trajectory can be computed given additional constraints from the IP. The set of initial parameters of the track candidate, referred to as tracklet parameters, is calculated using the full precision of the stubs. The tracklet parameters include: $\rho^{-1}$, the curvature of the trajectory; $t$, the tangent of the

dip angle, defined as $t = \sinh\eta$; $\phi_0$, the track direction in azimuthal angle at the point of closest approach to the IP; $z_0$, the $z$ coordinate of the track at the point of closest approach to the IP. For the seeding step, it is assumed temporarily that the tracking particle is produced promptly from the IP, therefore the transverse impact parameter $d_0$ is set to zero.

Given the computed parameters, a tracklet is then projected to other layers or disks to search for stubs that may belong to the same track. Projections to all possible barrel layers and disks are computed in parallel. The nominal projected position to a given layer (disk) is calculated with the average radius ($z$ coordinate) of the barrel layer (disk). Derivatives of $\phi$ and $z$ ($r$) in the barrel layer (disk) with respect to $r$ ($z$) are also computed, so that more precise projected coordinates can be obtained when comparing to a local stub from the difference of the stub's $r$ ($z$) and the nominal $r$ ($z$) of the layer (disk).

All stubs in the VM associated with the projection are considered for matching. Stubs within a predetermined search window around the projected position are accepted as candidates of matched stubs. The differences in $\phi$ and $z$ (or $r$ in disks) coordinates between these candidates and the projection are calculated. The stub that is the closest to the projection is selected as the matched stub and is associated to the seeding tracklet.

**Track fit**

A tracklet and all its matched stubs in other barrel layers and disks are used to compute the final track parameters. A tracklet has to have at least two matched stubs in order to be processed in this step, otherwise it is discarded. The track

parameters $\rho^{-1}$, $t$, $\phi_0$, $z_0$, and optionally $d_0$ are extracted from a linearized $\chi^2$ fit. They are determined by correcting the corresponding tracklet parameters calculated in the seeding step. The correction for each track parameter is obtained based on the $\phi$ and $z$ ($r$ for disks) residuals of the matched stubs and their derivatives with respect to the track parameter. In the barrel region, these derivatives to the first order only depend on the $r$ position and not on any track parameters. Therefore, these derivatives can be conveniently computed in advance and tabulated as lookup tables for FPGA implementations. For disks in endcap regions, these derivatives do have explicit dependence on the track parameter $t$ (i.e. $\sinh\eta$). However, given a set of matched stubs in disks, only a narrow range of $t$ (or equivalently $\eta$) is compatible with such hit pattern. As a result, a set of derivatives can be pre-computed for each hit pattern and stored in lookup tables. For a track candidate under consideration, the pattern of its associated stubs is used to look up the corresponding derivatives.

**Duplicate removal**

A same track can be reconstructed multiple times from different seeding combinations. In addition, fake tracks can arise from combinatorics. The duplicate removal step aims to keep only one reconstructed track for a tracking particle in the output track collection. Tracks found in different seedings are compared with each other. Two tracks that do not have at least three unique stubs are tagged as duplicates. The one with higher $\chi^2$ per degree of freedom, obtained previously from the fit, is discarded. Figure 6.9 shows distributions of number of reconstructed tracks per single muon event before and after the duplicate removal. As shown in the plot, more than 98% of total events have only one re-

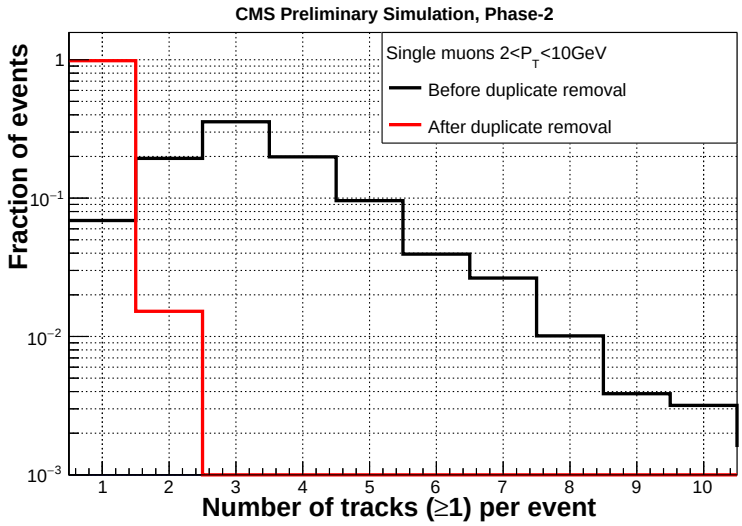


Figure 6.9: Distributions of number of tracks per single muon event before (black) and after (red) the duplicate removal. The plot is based on simulated events, each of which contains one muon with $p_{\mathrm{T}}$ between 2 and 10 GeV.

constructed track after the duplicate removal, which is expected for simulated single muon events.

### 6.3.2 Implementation

The Tracklet algorithm is implemented with integer calculation in both C++ emulation and firmware using Verilog hardware description language. The algorithm is also realized in C++ with floating point precision for validation and physics performance studies. The number of bits that are used to represent final track parameters as well as intermediate variables in the integer calculation are optimized for both sufficient precision and low resource usages on the FPGA. Resolutions of L1 track parameters calculated with the integer emulation are compared with resolutions of these parameters obtained based on floating point calculation in Fig. 6.10. As shown in these plots, the integer calculation barely degrades track parameter resolutions compared to the full precision floating point calculation.

The full algorithm is divided into several smaller processing steps, each of which is implemented as a type of processing modules. Each step generally contains multiple modules of the same type, and each of these parallel modules processes data in a particular detector sub-region. All processing modules are pipelined and are interfaced with either distributed RAMs (DRAM) or block RAMs (BRAM). An overview of the processing steps and memory types is given in Fig. 6.11. Processing steps and memory types are represented by red and blue boxes, respectively. Each box stands for a type of the module but not the actual instantiation. Grey arrows indicate the data dependencies among different steps. There are nine types of processing modules:

- **LayerRouter** (LR): sort and group input stubs based on their geometrical locations.

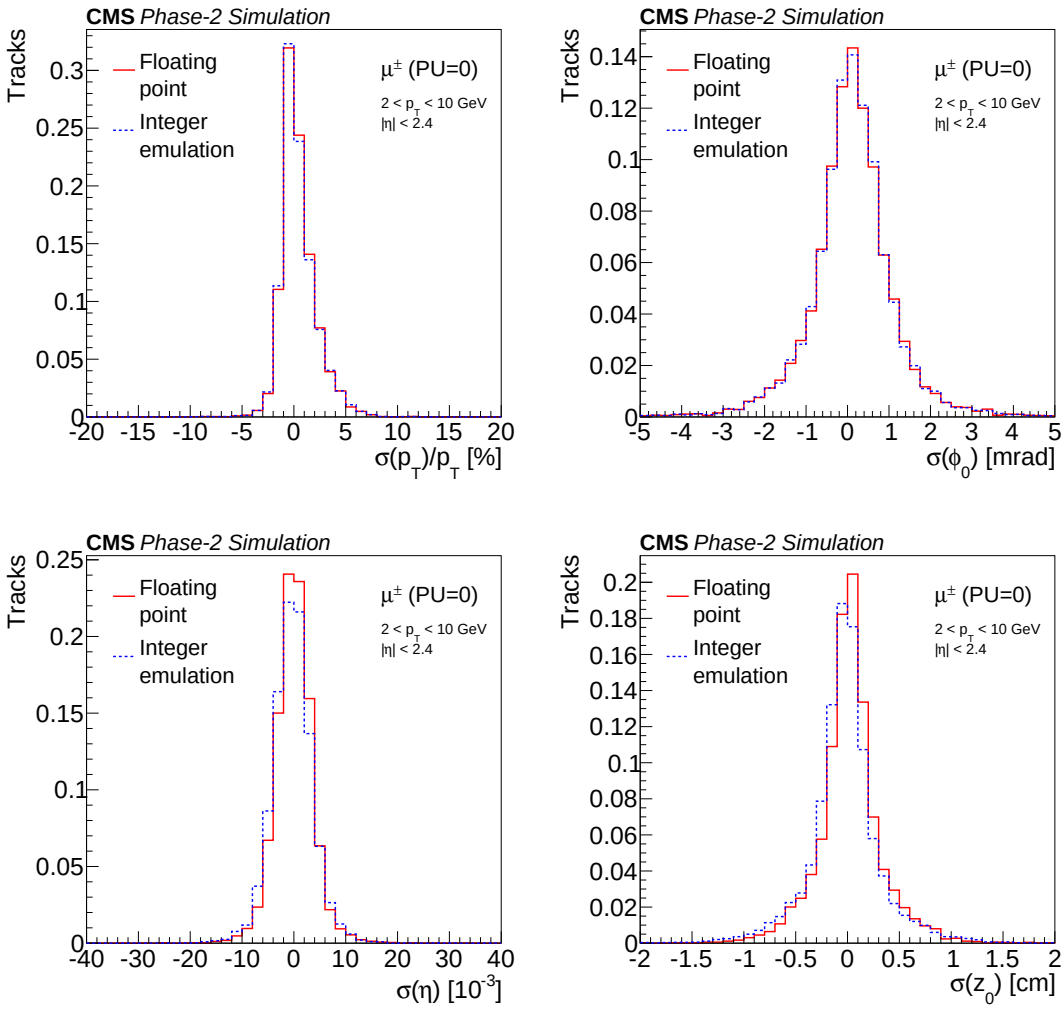


Figure 6.10: Resolutions of relative $p_{\mathrm{T}}$, $\phi_0$, $\eta$, and $z_0$ of L1 tracks obtained using full precision floating point calculation (solid red) and using the integer emulation (dashed blue). These plots are produced based on simulated tracks satisfying $2 < p_{\mathrm{T}} < 10$ GeV and $|\eta| < 2.4$ in single muon events without pileup.

- **VMRouter** (VMR): route stubs into their corresponding VMs in each layer/disk. The full precision stubs are stored in **AllStub** memory modules. Their indices and reduced stub information are stored in **VMStubs** memory modules. The reduced format in **VMStubs** saves resource usages and is sufficient for the purpose of forming seeds or finding matched stubs. The stub index carried by the VMStub can be used to access the

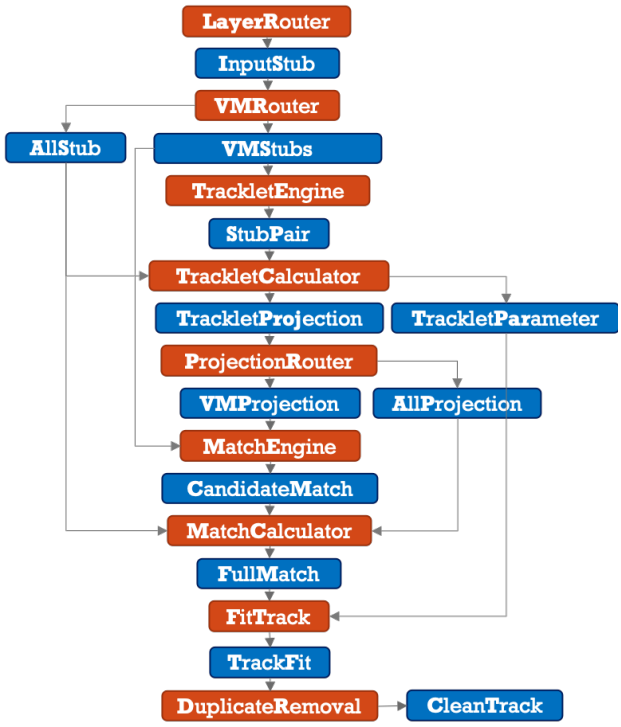


Figure 6.11: An overview of processing steps (red) and memory types (blue) of the project. Grey arrow indicate the data dependencies among the steps.

corresponding full precision stub if needed.

- **TrackletEngine** (TE): loop over stubs in an inner VM and an outer VM to form stub pairs as track seeds. As mention previously in Section 6.3.1, TrackletEngines are instantiated for only VM pairs that are compatible with tracks whose $p_T$'s are above the threshold. A TE uses the reduced stub format and stores the indices of stub pairs in the **StubPair** memory.
- **TrackletCalculator** (TC): read in stub pairs found by TEs, use the indices to look up the corresponding full precision stubs. The TC then calculates the initial tracklet parameters with the full stubs and stores the calculated parameters for each tracklet in the **TrackletParameter** memory. It also calculates projections to other layers/disks and stores them in **TrackletProjection** memories.

- **ProjectionRouter** (PR): route projections calculated by TCs in each layer/disk to their corresponding VMs. Projections with the reduced format are stored in **VMProjection** memories. They also carry indices that can point back to the full precision projections, which are stored in the **AllProjection** memories.
- **MatchEngine** (ME): loop over reduced stubs in the **VMStubs** memory and reduced projections in the **VMProjection** memory to find matched stub-projection pairs. The indices of the stub and the projection in the pair are stored in the **CandidateMatch** memory.
- **MatchCalculator** (MC): access the full precision stub and projection using indices stored in the matched pair, and calculate residuals between the stub and the projection. The MC then stores residuals to their corresponding **FullMatch** memories based on the seed of the projection.
- **FitTrack** (FT): perform linearized $\chi^2$ fit to estimate track parameters. Each FT module handles the track fit for one seeding combination. The fit uses initial tracklet parameters stored in **TrackletParameter** memories and updates them based on residuals stored in **FullMatch** memories. The final track parameters as well as the indices of stubs that are associated to the track are stored in the **TrackFit** memory.
- **DuplicateRemoval** (DR): remove duplicate tracks given reconstructed tracks in all seeding combinations. Tracks after duplicate removal are stored in the **CleanTrack** memory, which will then be sent out to downstream systems.

Each processing module starts with input memories filled by the previous step and processes data with a fixed amount of time per event before moving

on to the next event. Memory modules are all paginated, with each page assigned to one event. The data always carry an event identifier, which is used by processing modules as the top bits of memory read/write addresses. Doing so ensures that they always access the memory page corresponding to the right event. Most memory modules only needs one bit from the event identifier and basically function as ping-pong buffers. For **AllStub**, **AllProjection**, and **TrackletParameter** memories, because they are accessed more than one steps later, as shown in Fig. 6.11, more bits from the event identifier for read/write addresses, thus deeper memories, are needed.

Because each step has a fixed amount of time to process one event, the latency of the full system is also fixed. In case there are too much input data for a module to process in a given event, the extra data are ignored, leading to an algorithmic inefficiency due to the truncation. The emulation can mimic the truncation effect in hardware by using buffers with finite depths. Project configurations, including virtual module definitions and memory depths, are studied in the emulation and are optimized to mitigate such truncation effect. The redundancy provided by multiple seeding combinations also helps to recover tracks that may otherwise be lost due to truncation.

The full system contains about 1000 processing module instances and about 4000 memory modules. The instantiation of all these modules and their connections in the top-level firmware module is handled automatically by a Python script. A graph data structure is used to represent the project configuration. A partial project involving only a few steps or a specified module can also be generated from the script for testing purposes.

### 6.3.3 Hardware demonstrator

A version of the Tracklet project is implemented on a hardware demonstrator. The demonstrator aims to show the feasibility and performance of the Tracklet approach by implementing an end-to-end system on the currently available hardware. The hardware used for the demonstrator are the so-called CTP7 boards [102], which were developed for the current CMS Level-1 calorimeter trigger. The CTP7 is a $\mu$TCA board containing a Xilinx Virtex-7 (XC7VX690T) FPGA [103] and a Xilinx Zynq-7000 SoC.

The version of the firmware implemented on the demonstrator is not exactly the same as the latest version. Instead of the hourglass configuration, the $\phi$ sector definition with straight boundaries was adopted in the demonstrator firmware. This old sector definition is indicated by gray lines in Fig. 6.6. Given such $\phi$ sectors, a track with $p_{\mathrm{T}} > 2$ GeV can span more than one but at most two sectors. This means that each sector now needs to communicate with its two neighboring sectors regarding projections and matched stubs. For this purpose, two types of processing steps that handle inter-sector data transmissions are added to the project. Some memories that store projections and residuals of matched stubs also have to be deeper to account for the longer latency due to the inter-sector data transmission. In addition, the tracker layout assumed for the demonstrator was also slightly different from the tilted geometry. A flat geometry, as shown in Fig. 6.13, was used for the tracker, in which all barrel modules are parallel to the beam axis. Except for the necessity of neighboring sector data transmission, these differences do not affect the main algorithm beyond minor adjustments in some calculations.

A total of four CTP7 boards are used to build the demonstrator. As illus-

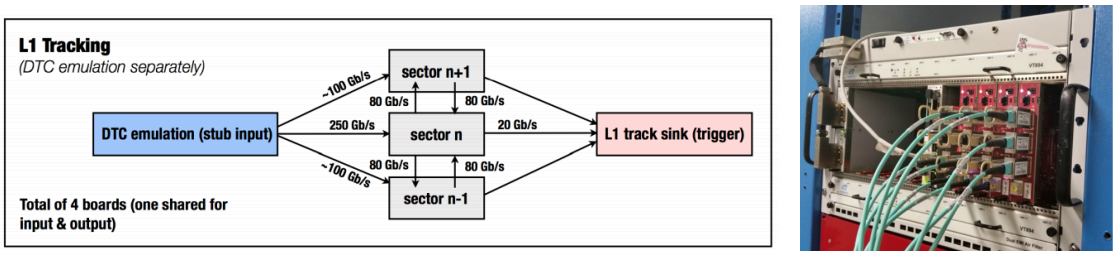


Figure 6.12: Left: a schematic overview of the Tracklet demonstrator. Three CTP7 boards are used as sector processing boards, and one CTP7 board is used as both the input stub source and as the output track sink. Estimated data transmission rates on the inter-board links are indicated as well. Right: a photo of the demonstrator system in a $\mu$CTA crate.

trated in Fig. 6.12, three CTP7 boards are used as sector processors: one for the central $\phi$ sector, two for its $\pm\phi$ neighboring sectors. The fourth CTP7 board functions as both an input emulator and an L1 track sink. Input stub data generated from simulation are sent to block RAMs in the FPGA of this board via the Zynq SoC. They are then distributed to the three sector processors via optical links. Reconstructed L1 tracks from all sector processors are sent back to this board and are stored in the block RAMs allocated as track sinks. These tracks are then read out by the Zynq SoC for validations and performance studies. In addition to the four CTP7 boards, an AMC13 card [104] is integrated into the demonstrator system to distribute a global 240 MHz clock signal to the CTP7 boards. This clock signal synchronizes all sector boards and is used as their processing clocks. Optical links with 10 Gbps speed and 8b/10b encoding are used for inter-board data transmissions.

The reported FPGA resource usage by the Xilinx Vivado software for a full-sector project with all seeding combinations is shown in Table 6.1. Fractions of needed resources over what are available on the Virtex-7 (690T) FPGA, as well as on other more powerful Virtex UltraScale+ FPGAs, are also shown in the table. The Virtex-7 FPGA does not have enough resources, particularly BRAMs,

Table 6.1: FPGA resource usages for a full-sector project report by Vivado (top row). Fractions of these resources over what are available on different FPGAs, including the Virtex-7 used for the demonstrator, are shown in the following rows.

| | LUT Logic | LUT Memory | BRAM | DSP |
|---|---|---|---|---|
| Full sector | 279733 | 151191 | 2721.5 | 1818 |
| Virtex-7 (690T) | 65% | 87% | 185% | 51% |
| VU3P | 32% | 81% | 85% | 80% |
| VU5P | 21% | 53% | 58% | 52% |
| VU7P | 16% | 40% | 42% | 40% |
| VU9P | 11% | 27% | 28% | 27% |
| VU11P | 10% | 27% | 29% | 20% |
| VU13P | 7% | 20% | 22% | 15% |

to implement a full-sector project. In order to demonstrate the feasibility to cover the full $\eta$ range of the detector, two partial projects are generated and implemented on the demonstrator: one for a half barrel ($+z$) region, referred to as a barrel-only project, and the other for quarter of the barrel plus the forward endcap region, referred to as the hybrid-region project. The barrel-only project considers tracks that only traverse the barrel region of the tracker. Only seedings with barrel layers are implemented in this project. The hybrid-region project covers tracks that traverse the endcap disk region of the detector as well as tracks in the transition region between barrel and endcaps. The tracker regions covered by the barrel-only and hybrid-region projects are indicated by solid blue and dashed green squares in Fig. 6.13, respectively.

For the demonstrator, a time multiplexing factor of six is adopted, which means each step of the system has $6 \times 25$ ns $= 150$ ns per event to process input data and fill its output memories. The latency of each processing step is the sum of the time needed to get the first output, typically 1-50 clock cycles, and the 150 ns processing time for each event. Table 6.2 lists the estimated latency of each processing step and the total system. The clock frequency is 240 MHz.

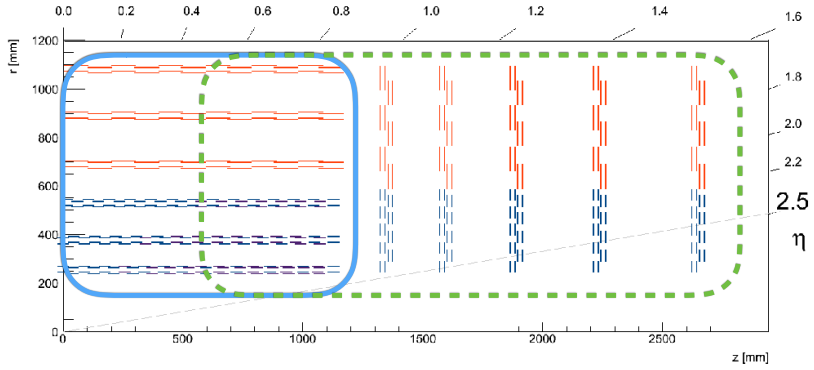


Figure 6.13: Tracker regions covered by the barrel-only (solid blue) and hybrid-region (dashed green) projects for the demonstrator. The flat tracker geometry shown here is the layout assumed by the version of the Tracklet firmware that was built for the demonstrator.

For steps that handle data transmissions with neighboring sectors (Projection Transceiver and Match Transceiver), an additional latency due to inter-board communications is included. The link latency is measured to be 316.7 ns based on 10 Gbps links with 8b/10b encoding. It is worth mentioning that the optical fibers used in the demonstrator are about 15 meters long. The total latency of the system between receiving the first input stub and sending out the first track (first-in first-out) is estimated to be 3345.8 ns. The first-in last-out latency is this value plus 150 ns processing time. The total latency is also measured with a clock counter implemented in the FPGA of the fourth CTP7 board. The measured first-in first-out latency is 800 clock cycles with a 240 MHz clock, corresponding to 3333 ns. The measured latency agrees well with the estimated one based on the latency model, both of which meet the 4 $\mu$s latency goal for L1 tracking. With the new hourglass configuration, the data transmission steps are not needed anymore. In this case, the total latency is largely reduced. Improvements in link protocol and higher link speed also help to further reduce the total latency.

Table 6.2: The demonstrator latency model. Estimated latencies for each processing step and for the total system are shown.

| Step | Proc. time (ns) | Step latency (CLK) | Step latency (ns) | Link delay (ns) | Step total (ns) |
|---|---|---|---|---|---|
| Input link | 0.0 | 1 | 4.2 | 316.7 | 320.8 |
| Layer Router | 150.0 | 1 | 4.2 | - | 154.2 |
| VM Router | 150.0 | 4 | 16.7 | - | 166.7 |
| Tracklet Engine | 150.0 | 5 | 20.8 | - | 170.8 |
| Tracklet Calculation | 150.0 | 43 | 179.2 | - | 329.2 |
| Projection Trans. | 150.0 | 13 | 54.2 | 316.7 | 520.8 |
| Projection Router | 150.0 | 5 | 20.8 | - | 170.8 |
| Match Engine | 150.0 | 6 | 25.0 | - | 175.0 |
| Match Calculator | 150.0 | 16 | 66.7 | - | 216.7 |
| Match Trans. | 150.0 | 12 | 50.0 | 316.7 | 516.7 |
| Track Fit | 150.0 | 26 | 108.3 | - | 258.3 |
| Duplicate Removal | 0.0 | 6 | 25.0 | - | 25.0 |
| Output Link | 0.0 | 1 | 4.2 | 316.7 | 320.8 |
| Total | 1500.0 | 139 | 579.2 | 1266.7 | 3345.8 |

Stub input data are generated from both simulated single muon events and simulated $t\bar{t}$ events with 200 pileup. Reconstructed track parameters from the demonstrator using these stub inputs are compared to the expected tracks computed with the integer-based C++ emulation. For single muon events, the demonstrator result agrees 100% with the emulation expectation. In busier $t\bar{t}$ events with 200 pileup, the agreement on final track parameters between the demonstrator and the emulation is better than 99%. Figure 6.14 shows very good agreements between track parameters computed by the emulation and those by the firmware simulation using simulated $t\bar{t}$ events with 200 pileup. The emulation and firmware results are indicated by solid black and dashed red histograms, respectively.

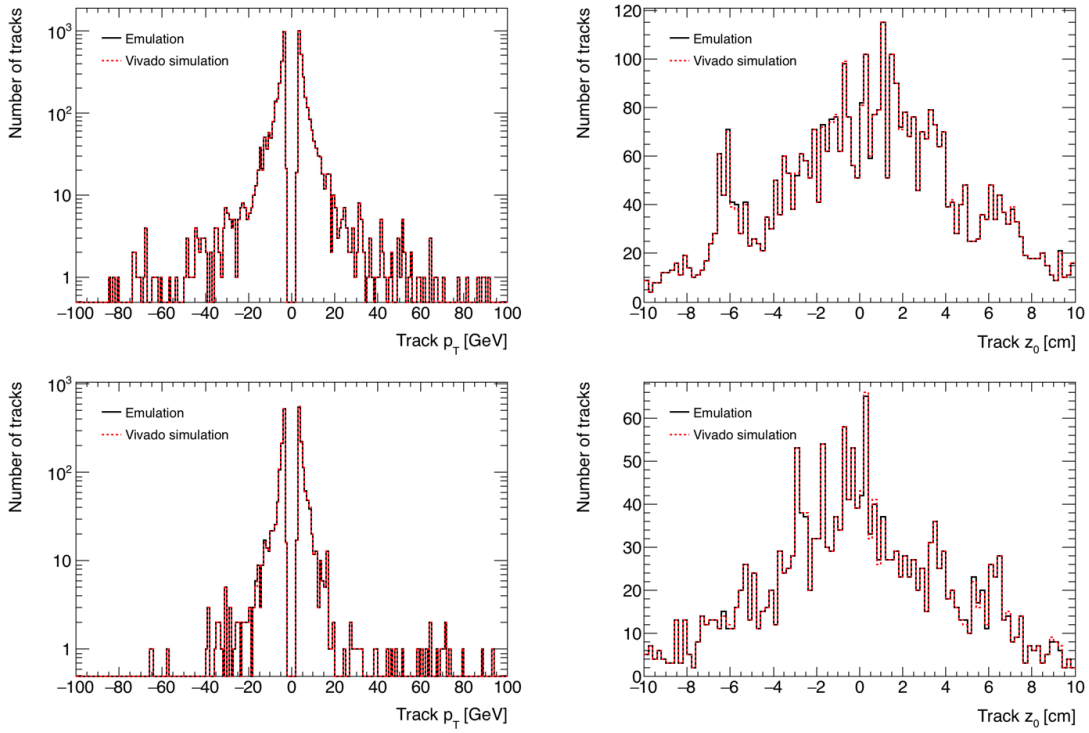


Figure 6.14: Comparisons of final track parameters $p_T$ (left) and $z_0$ (right) calculated in emulation (black) and in firmware simulation (red) in simulated $t\bar{t}$ events with an average pileup of 200. The barrel-only project plots are shown in the top row, and plots of the hybrid-region project are shown in the bottom.

### 6.3.4 L1 tracking performance

The perfect agreement in track parameters between the integer-based C++ emulation and the hardware demonstrator allows us to study the performance of the L1 tracking system confidently in the emulation. This section shows the L1 tracking performance of the latest project configuration with the hourglass sector definition and the tilted tracker geometry. The tracking efficiency and resolutions of track parameters are studied with the emulation under several pileup scenarios. The efficiency to find charged tracks in simulated $t\bar{t}$ events as a function of $p_T$ and $\eta$ is shown in Fig. 6.15. The efficiency is plotted separately for tracks in low $p_T$ range of $2 < p_T < 8$ GeV and in high $p_T$ range

$8 < p_T < 100$ GeV. Overall good tracking efficiencies (> 90%) are achieved in the busy $t\bar{t}$ events.

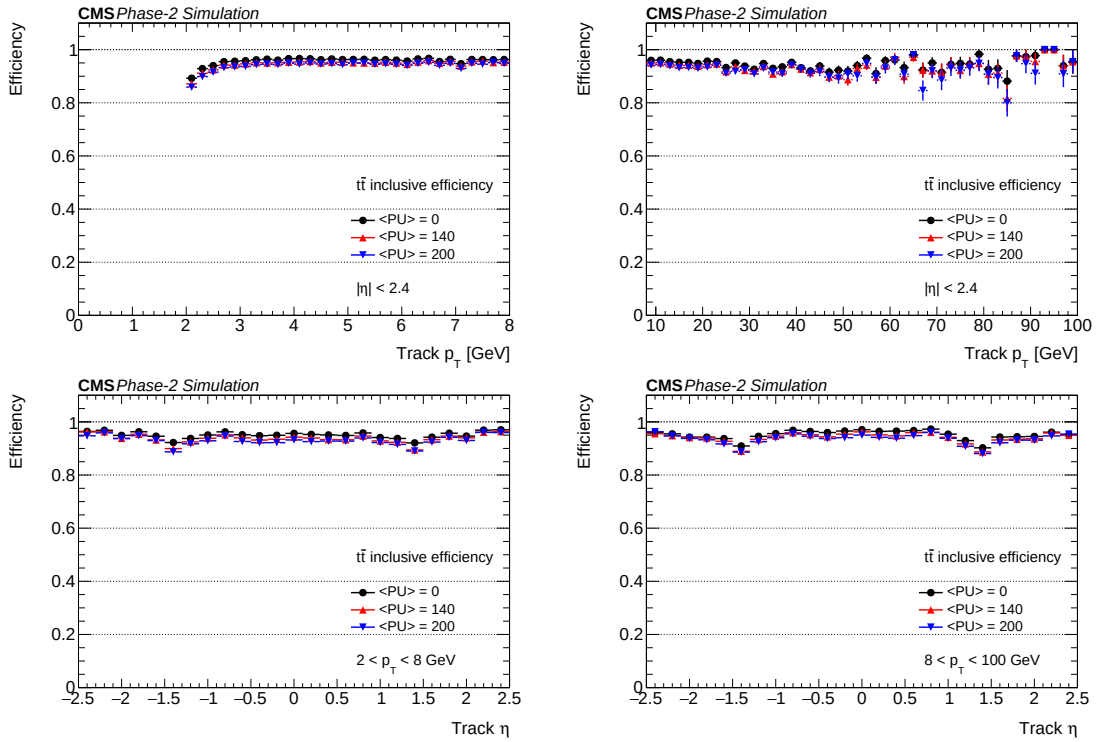


Figure 6.15: Tracking efficiency in simulated $t\bar{t}$ events with an average pileup of 0 (black), 140 (red), and 200 (blue). Efficiencies as a function of track $p_T$ and $\eta$ are shown in the top and bottom row, respectively. Plots in the left column are for low $p_T$ range ($2 < p_T < 8$ GeV), and those on the right are for high $p_T$ range ($8 < p_T < 100$ GeV).

Track parameter resolutions for relative $p_T$, $\phi_0$, $\eta$, and $z_0$ as a function of $|\eta|$ in simulated $t\bar{t}$ events with an average pileup of 200 are shown in Fig. 6.16. The resolutions are shown separately for low and high $p_T$ ranges. Good angular resolutions are shown in $\phi_0$ and $\eta$ resolution plots. The relative $p_T$ resolution varies from about 1% in the central barrel region to 4% in the endcap regions. The worse $p_T$ resolution at larger $|\eta|$ is due to longer extrapolation distances to the interaction point from endcap disks. Tracks in the low $p_T$ range have better $p_T$ resolutions than those with higher $p_T$ as expected. The $z_0$ resolution is about 1

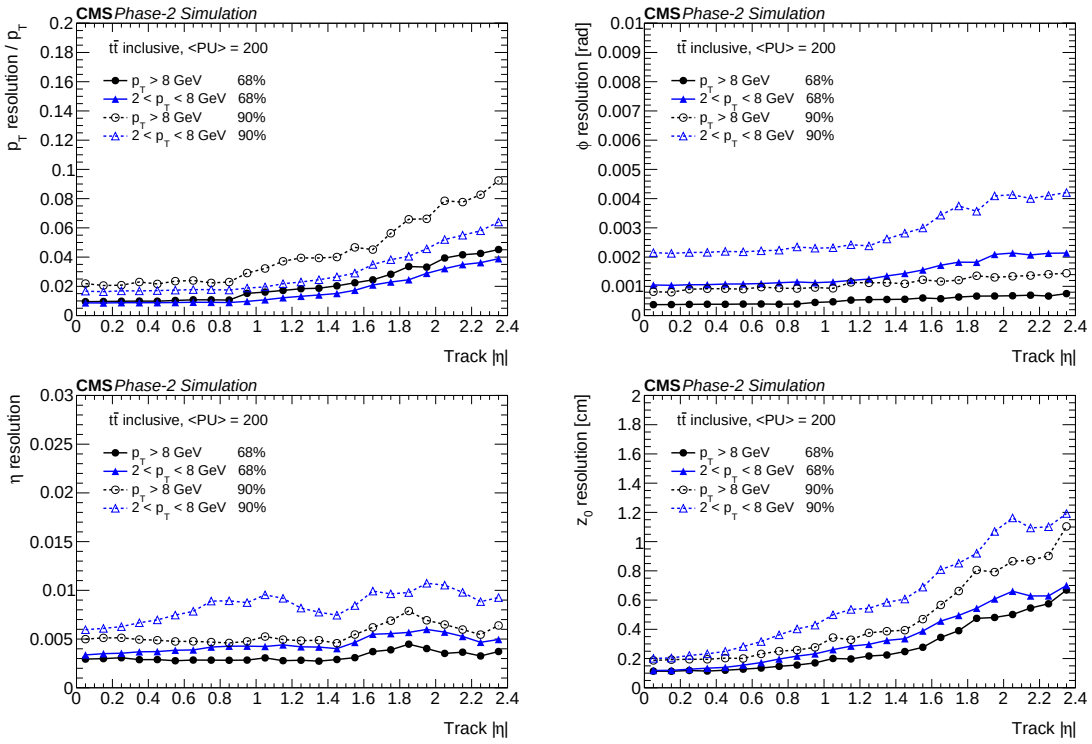


Figure 6.16: Resolutions of final track parameters $p_{\mathrm{T}}$ (top left), $\phi_0$ (top right), $\eta$ (bottom left), and $z_0$ (bottom right) as a function of track $|\eta|$. Simulated $\mathrm{t\bar{t}}$ events with an average of 200 pileup are used for the plots. Low $p_{\mathrm{T}}$ range ($2 < p_{\mathrm{T}} < 8$ GeV) and high $p_{\mathrm{T}}$ range ($8 < p_{\mathrm{T}} < 100$ GeV) are indicated by blue and black markers, respectively. Resolutions shown in these plots correspond to intervals that encompass 68% (solid marker) and 90% (open marker) of tracks in the corresponding track parameter distribution.

mm in the central barrel region and increases to about 6 mm in the most forward region considered by L1 tracking. The good $z_0$ resolution is sufficient for the L1 trigger to distinguish pileup vertices.

## 6.4 Summary

The High-Luminosity LHC is expected to start colliding proton bunches in 2026 at $\sqrt{s} = 14$ TeV with an average number of pileup up to 200. In order to handle such high pileup condition at the HL-LHC, CMS needs tracking information in the L1 trigger from its upgraded silicon strip tracker. The system that provides tracks to the L1 trigger is required to process input data at 40 MHz collision rate within 4$\mu$s latency budget.

The proposed Tracklet approach addresses the L1 tracking challenge via massive parallelization. The parallelization is achieve both in time via time-multiplexing and in space by partitioning and connecting detector sub-regions. The Tracklet approach implements a road search algorithm on an all-FPGA system. Track seeds, i.e. tracklets, are formed from a pair of stubs in adjacent layers/disks and are projected to other layers/disks to search for matched stubs. The final track parameters are extracted from the seed and all its matched stubs via a linearized $\chi^2$ fit. Duplicate tracks from multiple seeding combinations are removed before L1 tracks are sent out to downstream.

The algorithm is implemented in firmware using integer-based calculations. While DSPs are largely used to handle computations, many complicated calculations in the Tracklet algorithm are tabulated in LUTs. A version of the Tracklet firmware is successfully implemented on a hardware demonstrator with Xilinx Virtex-7 FPGAs. The demonstrator showed the feasibility and good performance of the Tracklet approach and convinced CMS to pursue an FPGA-based approach for its future L1 tracking system.

Developments on L1 tracking system are actively ongoing. Both hardware

and algorithms have been evolving and improving over the years. This section aims to provide a snapshot of the work being done in the Tracklet approach. Meanwhile, new hardware prototypes with more powerful FPGAs and faster optical links have been developed and tested. The feasibility to extend the seeding strategy to incorporate displaced tracks in addition to the prompt ones has been under study. Efforts have also been made to combine the strength of the Tracklet and TMTT approaches to define a so-call "hybrid" reference algorithm. Detailed discussions on these latest and ongoing developments are out of scope of this thesis.

# CHAPTER 7

# CONCLUSION

This thesis presents the measurement of $t\bar{t}H$ cross section in final states with electrons, muons, and hadronically decaying taus in proton-proton collisions at a center-of-mass energy of 13 TeV with the CMS experiment. This analysis channel targets the Higgs decay mode $H \rightarrow WW$, ZZ, and $\tau\tau$. Events are selected and grouped into categories based on multiplicities of leptons and hadronic taus. Background processes are estimated based on both Monte Carlo simulations and recorded data in control regions. Dedicated multivariate discriminants are developed for object identifications and signal extractions. The $t\bar{t}H$ signal strength is extracted from a maximum likelihood fit based on the distribution of event-level multivariate discriminants.

The analysis strategy discussed in this thesis is based on that applied on data recorded in the year of 2017, which corresponds to an integrated luminosity of 41.5 $fb^{-1}$. A similar but older version of the analysis was performed on data recorded in 2016, corresponding to an integrated luminosity of 35.9 $fb^{-1}$. A statistical combination of results in both years is performed. The measured $t\bar{t}H$ signal strength, based on 35.9 + 41.5 $fb^{-1}$ of data, is $\mu = 0.96^{+0.34}_{-0.31}$. The measured result is in good agreement with the SM expectation within its uncertainties.

A proposed approach to provide track information to the future CMS Level-1trigger is also presented. In the era of HL-LHC, CMS needs tracks at Level-1 trigger to handle the high pileup condition. The proposed Tracklet approach can handle input data rate at 40 MHz and reconstruct tracks within 4 $\mu$s. It adopts a road search algorithm and implements a highly parallel system on the FPGA. A version of the Tracklet firmware was realized on a demonstrator using currently

available hardware, showing both the feasibility and good performance of this approach for Level-1 tracking.

APPENDIX A

# AUXILIARY MEASUREMENTS FOR REDUCIBLE BACKGROUND ESTIMATIONS

## A.1 Lepton fake factor

The lepton fake factor is defined as the probability with which a non-prompt or misidentified lepton that has satisfied the fakeable selection criteria also passes the tight selection. The fakeable and tight selection criteria are defined in Table 5.3 and 5.4 for muons and electrons, respectively. Measurements of fake factors are carried out separately for electrons and muons in a so-called determination region (DR) containing mostly events with multiple jets produced from QCD processes. Events selected in the DR are required to have exactly one fakeable lepton and at least one jet that is well separated from the lepton by $\Delta R > 0.7$. Events in data are recorded using a set of single lepton and single lepton plus a PF jet triggers, as listed in Table A.1. No isolation condition is required by any of the triggers. Depending on which HLT path triggers the event, the reconstructed lepton and jet $p_{\mathrm{T}}$'s are required to be greater than the thresholds listed in the table in order to ensure that the trigger is fully efficient. In addition, for a given trigger path, only events with lepton cone-$p_{\mathrm{T}}$ in the range specified in the table are selected in the DR for measuring the fake factor.

### A.1.1 Fakeable lepton definition tuning

The success of the Fake Factor method for estimating the non-prompt lepton background relies on the assumption that lepton fake factors measured in the

Table A.1: Triggers used to record events for the measurement of the lepton misidentification rate.

| HLT path | Lepton cone-$p_T$ | Lepton $p_T$ | Jet $p_T$ |
|---|---|---|---|
| HLT_Ele8_CaloIdM_TrackIdM_PFJet30 | 15-45 GeV | > 27 GeV | > 30 GeV |
| HLT_Ele17_CaloIdM_TrackIdM_PFJet30 | 25-100 GeV | > 17 GeV | > 30 GeV |
| HLT_Ele23_CaloIdM_TrackIdM_PFJet30 | 32-100 GeV | > 23 GeV | > 30 GeV |
| HLT_Mu3_PFJet40 | 10-32 GeV | > 3 GeV | > 45 GeV |
| HLT_Mu8 | 15-45 GeV | > 8 GeV | > 30 GeV |
| HLT_Mu17 | 32-100 GeV | > 17 GeV | > 30 GeV |
| HLT_Mu20 | 32-100 GeV | > 20 GeV | > 30 GeV |
| HLT_Mu27 | 45-100 GeV | > 27 GeV | > 30 GeV |

DR are the same as in the application region (AR). This cannot be taken as granted, especially the DR is populated with QCD multi-jet events, while the AR is enriched in $t\bar{t}$ samples containing heavy-flavored jets with potentially different kinematic distributions.[1] In particular, the isolation condition for a non-prompt lepton is sensitive to the $p_T$ and flavor of its mother parton, and so is the fake factor. The criteria of the fakeable lepton selection are carefully tuned to avoid biases in the fake factor measurement and to improve the closure of the non-prompt lepton background estimation.

In order to account for the effect of mother parton $p_T$, the fake factor is parameterized as a function of cone-$p_T$ instead of $p_T$. The lepton cone-$p_T$, defined in Section 5.2.1, corrects the reconstructed lepton $p_T$ by accounting for the energy in the isolation cone. As a result, the cone-$p_T$ of a non-prompt lepton provides a better estimation of the mother parton $p_T$. By defining the fakeable lepton collection using cone-$p_T$ and parameterizing the fake factor in bins of cone-$p_T$ (i.e. approximately $p_T$ of the mother parton), the difference of mother parton kinematics in the DR and the AR becomes irrelevant.

[1] While it may be possible to define an alternative $t\bar{t}$-enriched DR such that the DR and the AR are more similar, such DR would have less number of selected events and more prompt lepton contamination compared to the QCD multi-jet DR. Nevertheless, efforts are being made to measure lepton fake factors in the $t\bar{t}$-enriched DR for future version of the analysis.

While choices of identification criteria for fakeable leptons and the minimum prompt lepton BDT output required for tight leptons effectively determine the fake factor for non-prompt leptons from light-flavored jets, they do not affect much the fake factor for those from heavy-flavored jets. The requirement on the b-tagging discriminator of the matched jet (c.f. Table 5.3 and 5.4) is therefore introduced to provide an extra handle to tune fake factors for non-prompt leptons originated from heavy-flavored jets. Effects of these lepton selection criteria are studied in simulated samples enriched in either light- or heavy-flavored jets. By tuning these cuts for the fakeable lepton definition, the flavor dependency of lepton fake factors can be reduced. Small remaining discrepancies of fake factors in the DR and AR due to jet flavor composite are treated as systematic uncertainties.

Another potential bias of the fake factor measurement can be introduced by the different high-level triggers used in the DR (c.f. Table A.1) and in the AR (c.f. Table 5.7). Triggers used in the DR require the presence of either a lepton or a lepton plus a PF jet. No lepton isolation condition is required in these triggers. On the other hand, the AR uses single-, double-, or triple-lepton triggers with isolation requirements and with different $p_{\mathrm{T}}$ thresholds. In order to avoid such bias, lepton fake factors cannot depend on whether an event passing a certain trigger or not. This can be ensured if the fake factor measurement is performed in the region where the trigger is fully efficient. For this reason, a trigger with a given threshold is only used to measure fake factors for cone-$p_{\mathrm{T}}$ bins that are at least about twice of the trigger $p_{\mathrm{T}}$ threshold, as indicated in Table A.1. Although no lepton isolation is explicitly required in the triggers used in the DR, the electron identification criteria in the above electron triggers are not strictly looser than those applied in offline electron reconstructions. Such

bias is avoided by adding an additional set of selection criteria on $\sigma_{i\eta i\eta}$, H/E, 1/E - 1/p (c.f. Table 5.4) in fakeable and tight electron selections to mimic the identification requirements at trigger levels.

### A.1.2 Measurement of lepton fake factor

The lepton fake factors are measured separately for muons and electrons in bins of cone-$p_{\mathrm{T}}$ and $\eta$. In each bin, the fake factor is computed as

$$f_i = \frac{N_{\mathrm{pass}}}{N_{\mathrm{pass}} + N_{\mathrm{fail}}} \tag{A.1}$$

where $N_{\mathrm{pass}}$ ($N_{\mathrm{fail}}$) is the number of fakeable leptons in multi-jet events that pass (fail) the tight lepton selection. Events in each kinematic bin are categorized into "pass" and "fail" regions, in which $N_{\mathrm{pass}}$ and $N_{\mathrm{fail}}$ are determined together by a maximum likelihood (ML) fit. MC simulated samples are used to model contributions from non-prompt leptons and from prompt lepton contamination. Non-prompt leptons in multi-jet events are modeled using QCD MC samples. Prompt lepton contamination in the DR is mainly from W+jets events, with smaller contributions from Z+jets and diboson productions. These simulated events are summed up and labeled as "prompt background". Normalizations of simulated QCD multi-jet events and the prompt background are allowed to scale independently. A simultaneous ML fit to data in both pass and fail regions is then used to extract the two scale factors. With the fitted scale factor for QCD multi-jet events, $N_{\mathrm{pass}}$ and $N_{\mathrm{fail}}$, therefore $f_i$, are measured in the cone-$p_{\mathrm{T}}$ and $\eta$ bin.

The ML fit is performed on the distribution of an observable $m_{\mathrm{T}}^{\mathrm{fix}}$, defined as

$$m_{\mathrm{T}}^{\mathrm{fix}} = \sqrt{2 p_{\mathrm{T}}^{\mathrm{fix}} p_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos\Delta\phi)} \tag{A.2}$$

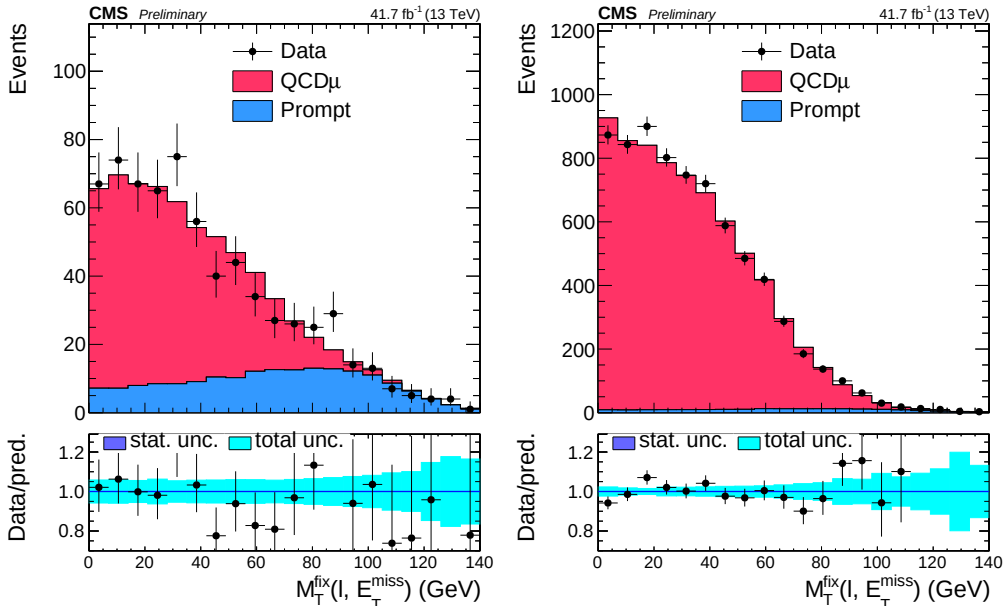


Figure A.1: Distributions of $m_{\mathrm{T}}^{\mathrm{fix}}$ in the pass (left) and fail (right) regions for muons satisfying 15 < cone-$p_{\mathrm{T}}$ < 20 GeV and $|\eta| > 1.2$.

where $p_{\mathrm{T}}^{\mathrm{fix}}$ is set to 35 GeV, $p_{\mathrm{T}}^{\mathrm{miss}}$ is the magnitude of the missing transverse momentum, and $\Delta\phi$ is the difference between the azimuthal angle of the lepton and that of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. This observable is inspired by the transverse mass variable used to measure the W boson mass. The only difference is that $p_{\mathrm{T}}^{\mathrm{fix}} = 35$ GeV is used instead of $p_{\mathrm{T}}$ of the lepton in order to make the observable lepton $p_{\mathrm{T}}$ independent. By doing so, correlations between the fitting observable and the lepton $p_{\mathrm{T}}$, thus the fake factor, are avoided. The observable $m_{\mathrm{T}}^{\mathrm{fix}}$ exploits the fact that its distribution in multi-jet events peaks at small value and is fast falling, since the missing energy in these events are usually small and are due to resolution effects, while $m_{\mathrm{T}}^{\mathrm{fix}}$ distribution in W+jets events has a broad peak around the W boson mass ($\sim$ 80 GeV). The lepton fake factor measurement is therefore improved due to the better separation between multi-jet events and the prompt background. Post-fit distributions of $m_{\mathrm{T}}^{\mathrm{fix}}$ in both pass and fail regions in some cone-$p_{\mathrm{T}}$ and $\eta$ bins are shown in Fig. A.1 and A.2 for muons and electrons, respectively.

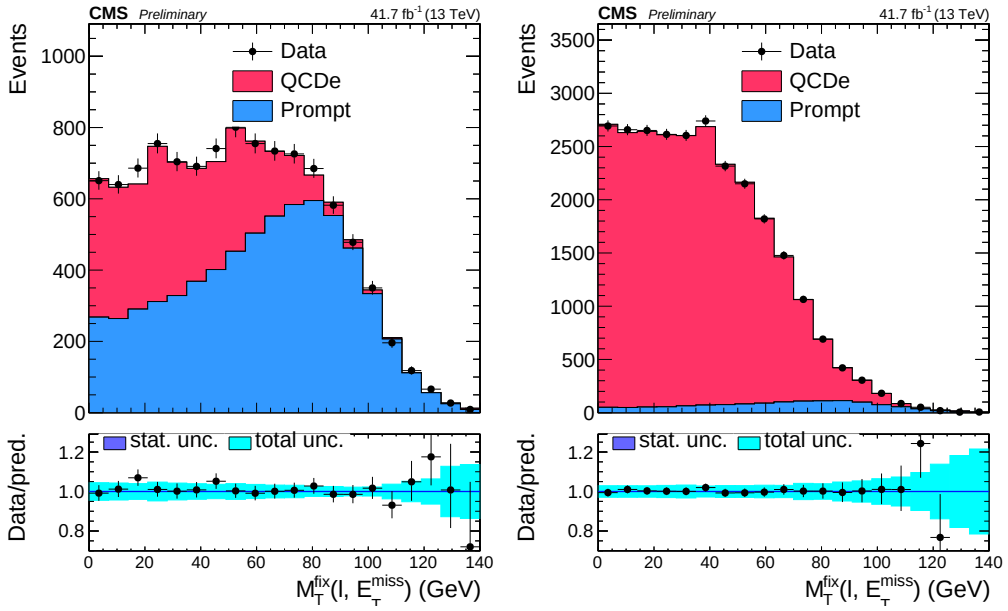


Figure A.2: Distributions of $m_{\mathrm{T}}^{\mathrm{fix}}$ in the pass (left) and fail (right) regions for electrons satisfying 25 < cone-$p_{\mathrm{T}}$ < 35 GeV and $|\eta| < 1.479$.

An additional correction is applied for electrons to account for the effect from photon conversion. The fraction of multi-jet events in which the selected fakeable electron is matched to a generator level prompt photon by $\Delta R < 0.3$ is determined in the simulation separately for the pass and fail regions in each cone-$p_{\mathrm{T}}$ and $\eta$ bin. $N_{\mathrm{pass}}$ and $N_{\mathrm{fail}}$ determined from the fit are scaled based on the corresponding fractions. Such correction reduces electron fake factors by about 10-20%.

The measured lepton fake factors in data for muons and electrons are shown in Fig. A.3 and A.4, respectively. Lepton fake factors extracted directly from $\bar{\mathrm{t}}\mathrm{t}$ and QCD MC samples are also included. For electrons, both fake factors without and with photon ($\gamma$) conversion corrections are shown in the plots.

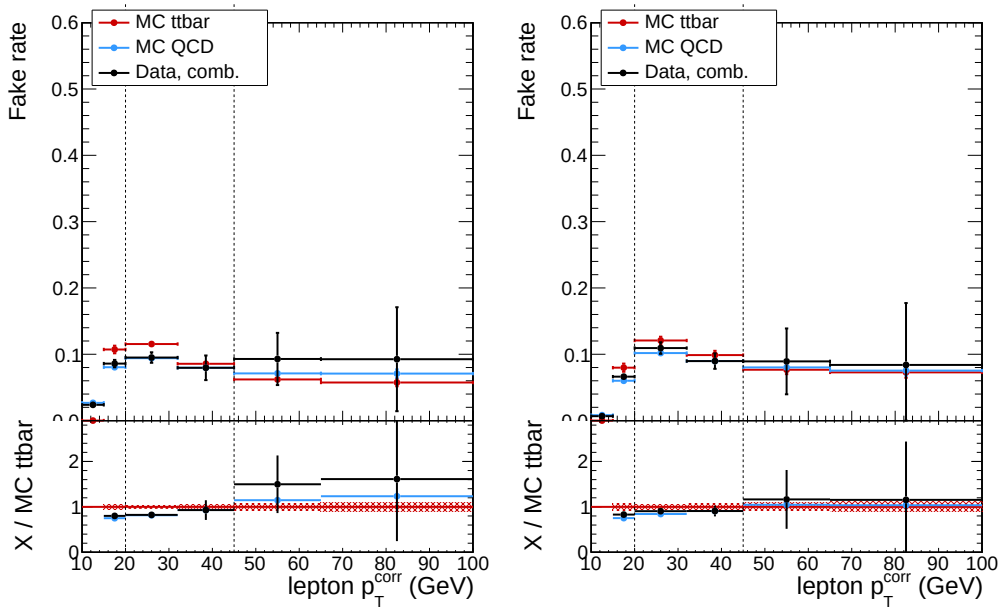


Figure A.3: Measured muon fake factors as functions of $p_T$ in the barrel (left) and endcap (right) regions.

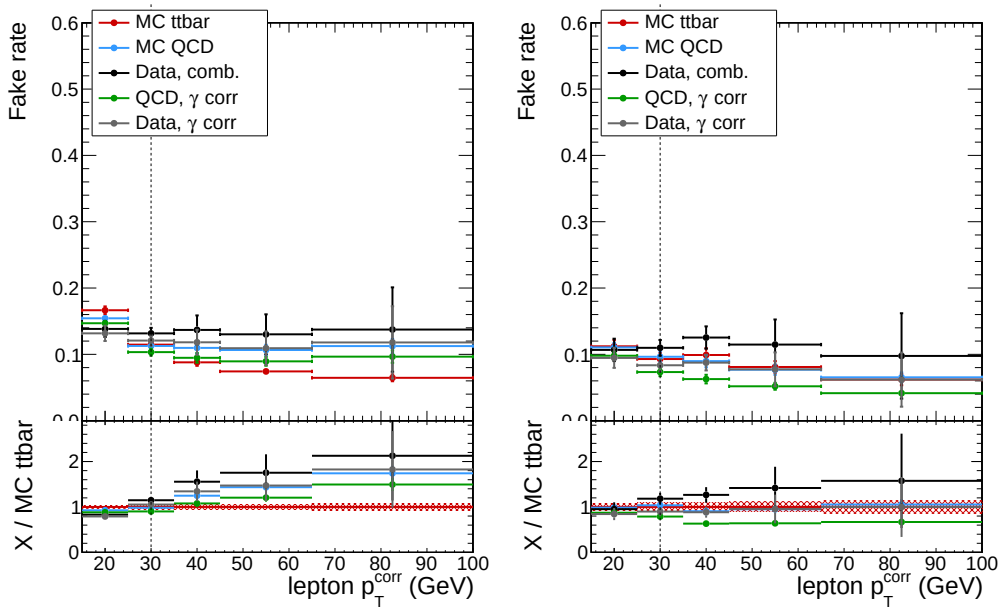


Figure A.4: Measured $\tau_h$ fake factors as functions of $p_T$ in the barrel (left) and endcap (right) regions.

## A.2 Hadronic tau fake factor

The fake factor for $\tau_h$ is defined as the probability with which a quark or gluon jet passes the tight $\tau_h$ selection given that it has passed the fakeable $\tau_h$ selection criteria (c.f. Table 5.6). The $\tau_h$ fake factor is measured using dileptonic $t\bar{t}$ + jets events in which one of the W boson decays into an electron and the other W decays into a muon.

Event selection criteria in the determination region for measuring $\tau_h$ fake factors are inspired by the $2\ell ss+1\tau_h$ signal region selection (c.f. Table 5.8). Events selected in the DR are required to contain one muon and one electron, both of which pass the tight selection criteria defined in Table 5.3 and 5.4, respectively. The two leptons are required to be of opposite charge. The leading lepton $p_T$ needs to be greater than 25 GeV, and the trailing one must have a $p_T$ no less than 15 GeV. At least one $\tau_h$ candidate passing the fakeable selection criteria (c.f. Table 5.6) is required. The selection criteria on jets, b tags, missing transverse momentum, and dilepton mass veto are the same as the $2\ell ss$ + $1\tau_h$ category, except the requirement on number of jets is relaxed to 2 instead of 3.

Data used in the $\tau_h$ fake factor measurement are recorded with single-lepton triggers or double-lepton triggers requiring an electron and a muon, as listed in Table 5.7. Using the same set of triggers in the DR and the corresponding AR ensures that the measured $\tau_h$ fake factors are unbiased and makes the measurement procedure relatively simpler. Simulated events are used to model dileptonic $t\bar{t}$ + jets process containing misidentified $\tau_h$, as well as contamination in the DR from background processes containing genuine $\tau_h$. All simulated events are required to pass the same set of trigger paths used to record data in the DR.

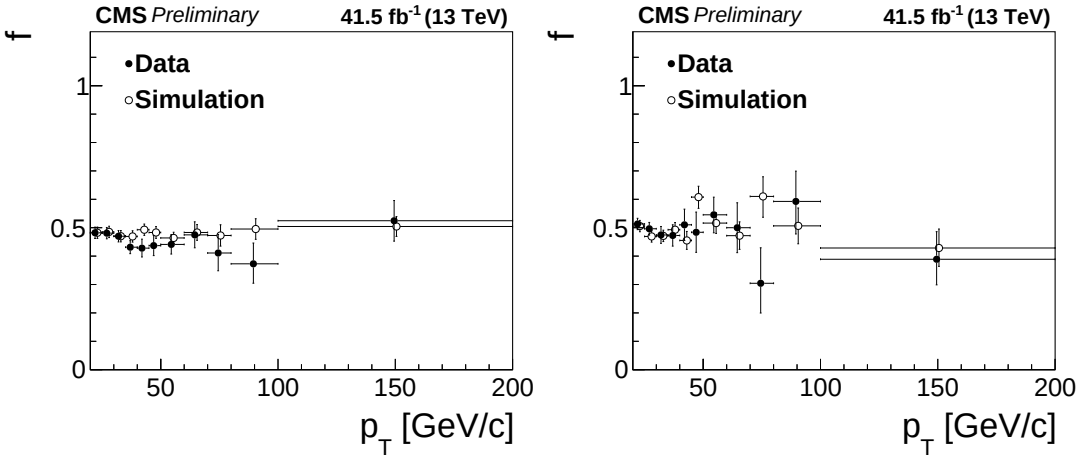


Figure A.5: Measured $\tau_h$ fake factors as functions of $p_T$ in the barrel (left) and endcap (right) regions. The tight $\tau_h$ definition for $2\ell ss+1\tau_h$ and $3\ell+1\tau_h$ categories with the Loose working point of $\tau_h$ MVA ID is used.

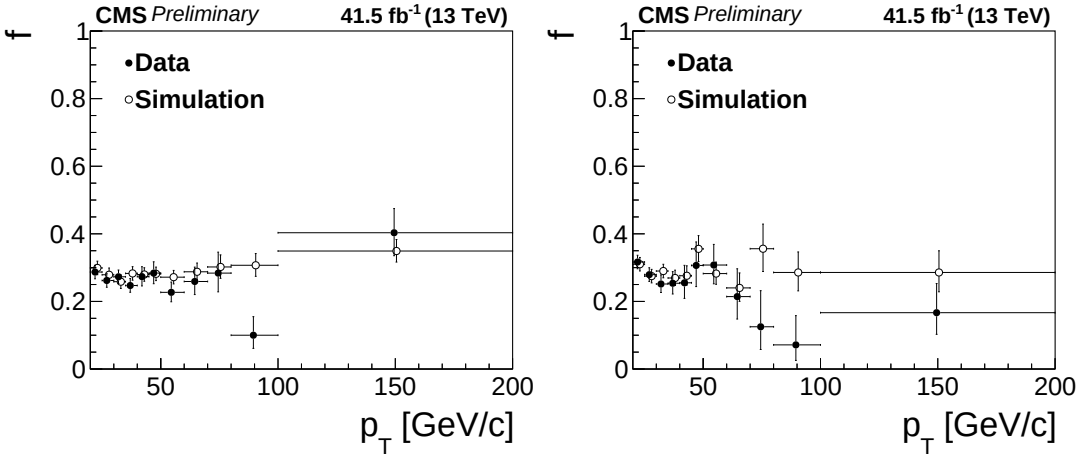


Figure A.6: Measured $\tau_h$ fake factors as functions of $p_T$ in the barrel (left) and endcap (right) regions. The tight $\tau_h$ definition for $1\ell+2\tau_h$ and $2\ell+2\tau_h$ categories with the Medium working point of $\tau_h$ MVA ID is used.

Measurements of $\tau_h$ fake factors are carried out in bins of $\tau_h$ $p_T$ and separately in the barrel ($|\eta| < 1.479$) and endcap ($|\eta| > 1.479$) regions. Due to different $\tau_h$ MVA ID working points used in the tight selection criteria (Loose WP for tight $\tau_h$ in $2\ell ss+1\tau_h$ and $3\ell+1\tau_h$ categories, Medium WP for tight $\tau_h$ in $1\ell+2\tau_h$ and $2\ell+2\tau_h$ categories), fake factor measurements are done for loose and medium $\tau_h$ MVA ID working points separately and are only used in the applicable categories. In

each $p_T$ and $\eta$ bin, the $\tau_h$ fake factor $f$ is computed as the ratio of the number of $\tau_h$ candidates that satisfy the tight selection criteria to the total number of $\tau_h$ candidates passing the fakeable selection. This calculation is done separately using data and MC simulated samples. The measured $\tau_h$ fake factors in data and simulation are shown in Fig. A.5 and A.6 for loose and medium WP of $\tau_h$ MVA ID, respectively, as functions of $p_T$ separately in the barrel and endcap regions.

The $\tau_h$ fake factor measured in data is not directly used in the fake factor method to estimate the misidentified $\tau_h$ background for the purpose of reducing the systematic uncertainty due to statistical fluctuation. Instead, the ratio of the fake factor measured in data to that in simulation as a function of $p_T$ in each $\eta$ region is fitted by a linear function $c_0 + c_1 p_T$. The fake factor used in the background estimation is obtained from multiplying the fake factor measured in simulation by the corresponding data-to-MC ratio from the fit. The systematic uncertainty on $\tau_h$ fake factors are estimated by varying the two fit parameters $c_0$ and $c_1$ based on their uncertainties. In order to account for the correlation between the two parameters, their covariance matrix $\mathrm{cov}(c_0, c_1)$ obtained from the fit is diagonalized to determine the eigenvectors and eigenvalues. The two parameters are then transformed based on the eigenvectors into two uncorrelated parameters, uncertainties of which are the corresponding eigenvalues. The data-to-MC ratio can be expressed in terms of the transformed parameters, and its uncertainties can be evaluated by varying the two uncorrelated parameters independently. The data-to-MC ratios and their fitted functions, as well as the variations based on eigenvalues of the covariance matrix are shown in Fig. A.7 and A.8 for loose and medium WP of $\tau_h$ MVA ID, respectively.

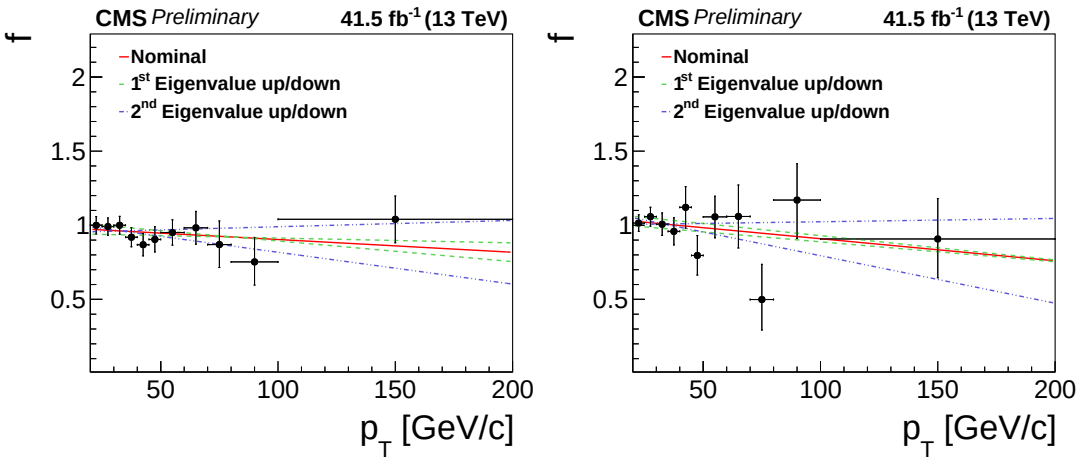


Figure A.7: Ratio of $\tau_h$ fake factors measured in data to those in MC simulation in the barrel (left) and endcap (right) regions. The ratio is fitted to a linear function $c_0 + c_1 p_T$. The tight $\tau_h$ definition for $2\ell ss + 1\tau_h$ and $3\ell + 1\tau_h$ categories with the Loose working point of $\tau_h$ MVA ID is used.

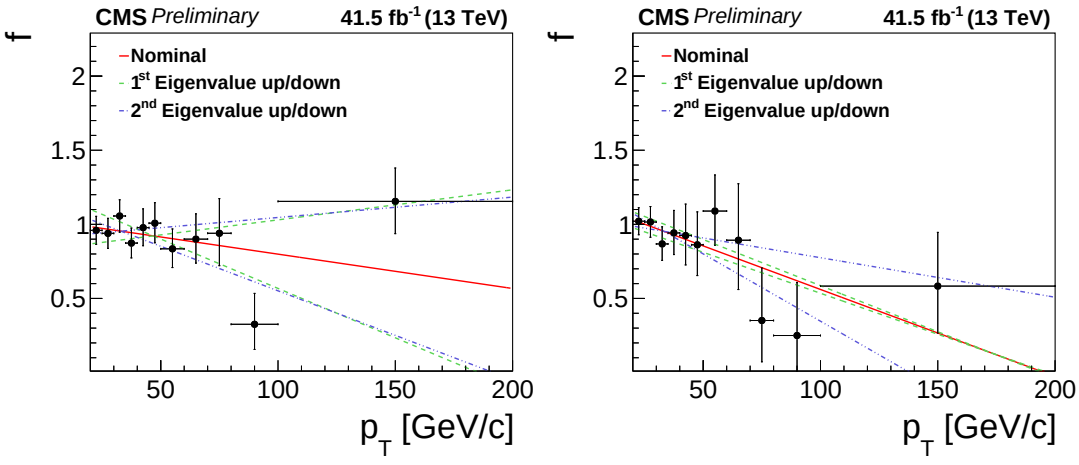


Figure A.8: Ratio of $\tau_h$ fake factors measured in data to those in MC simulation in the barrel (left) and endcap (right) regions. The ratio is fitted to a linear function $c_0 + c_1 p_T$. The tight $\tau_h$ definition for $1\ell + 2\tau_h$ and $2\ell + 2\tau_h$ categories with the Medium working point of $\tau_h$ MVA ID is used.

## A.3 Lepton charge misidentification rate

The probability to misidentify the charge of a lepton is determined in Drell-Yan events ($Z/\gamma^* \rightarrow \ell^+\ell^-$, $\ell = e, \mu$), in which the distribution of lepton pair masses peaks at the Z boson mass ($m_Z = 91.2$ GeV). The selected events for charge misidentification rate measurements are required to contain two electrons or muons satisfying the tight selection criteria defined in Table 5.4 or 5.3, respectively. Events are separated into the same-sign (SS) and opposite-sign (OS) regions based on charges of the two leptons. The following discussions focus on measuring electron charge misidentification rate. The muon charge misidentification rate can be measured in the same way and is found to be negligible after applying charge quality requirement in $2\ell$ss and $2\ell$ss + $1\tau_h$ categories (c.f. Section 5.3.4).

The electron charge misidentification rate $p$ is measured in three bins of $p_T$ and two bins of $\eta$. The three $p_T$ bins are labeled as "L" for the low $p_T$ region in the range $10 < p_T < 25$ GeV, "M" for the medium $p_T$ region with $25 < p_T < 50$ GeV, and "H" for the high $p_T$ region with $p_T$ greater than 50 GeV. Two $\eta$ bins are $|\eta| \leq 1.479$ for the barrel region, denoted as "B", and $1.479 < |\eta| < 2.5$ for the endcap region, denoted as "E". The 6 $p_T$-$\eta$ bins are thus labeled as "LB", "MB", "HB", "LE", "ME", and "HE".

For a pair of electrons, there are a total of $6^2$ bin combinations, 21 of which are unique. Since it is not known which one of the two electrons gets its charge mismeasured, the ratio of the event yield in the SS region ($N_{SS}$) to the total event yield ($N_{SS} + N_{OS}$) in each of the 21 categories is expressed as the sum of the two

Table A.2: Fractions of the same-sign events in each category expressed as sums of the corresponding charge misidentification rates.

| $r_{\text{bin1\_bin2}}$ | LB | MB | HB | LE | ME | HE |
|---|---|---|---|---|---|---|
| LB | $2p_{\text{LB}}$ | $p_{\text{LB}} + p_{\text{MB}}$ | $p_{\text{LB}} + p_{\text{HB}}$ | $p_{\text{LB}} + p_{\text{LE}}$ | $p_{\text{LB}} + p_{\text{ME}}$ | $p_{\text{LB}} + p_{\text{HE}}$ |
| MB | – | $2p_{\text{MB}}$ | $p_{\text{MB}} + p_{\text{HB}}$ | $p_{\text{MB}} + p_{\text{LE}}$ | $p_{\text{MB}} + p_{\text{ME}}$ | $p_{\text{MB}} + p_{\text{HE}}$ |
| HB | – | – | $2p_{\text{HB}}$ | $p_{\text{HB}} + p_{\text{LE}}$ | $p_{\text{HB}} + p_{\text{ME}}$ | $p_{\text{HB}} + p_{\text{HE}}$ |
| LE | – | – | – | $2p_{\text{LE}}$ | $p_{\text{LE}} + p_{\text{ME}}$ | $p_{\text{LE}} + p_{\text{HE}}$ |
| ME | – | – | – | – | $2p_{\text{ME}}$ | $p_{\text{ME}} + p_{\text{HE}}$ |
| HE | – | – | – | – | – | $2p_{\text{HE}}$ |

electrons' charge misidentification rates.[2] A complete list of these fractions $r$, expressed in terms of electron charge misidentification rates $p$, are shown in Table A.2.

The electron charge misidentification rates in the six bins can then be obtained by solving the over-constrained linear system if the fractions of events in the SS region are known. The event yields of Drell-Yan process in the SS and OS regions , thus the SS event fractions, in each of the 21 categories are determined by a maximum likelihood fit of the dielectron mass distribution in the range of $60 < m_{ee} < 120$ GeV. Shape templates of Drell-Yan and background processes are modeled using the MC simulation. The fit is done independently in each category, but simultaneously in the SS and OS regions to extract $N_{\text{SS}}$ and $N_{\text{OS}}$. The measured electron charge misidentification rates in the six $p_{\text{T}}$-$\eta$ bins are summarized in Table A.3 for data recorded in 2017.

Table A.3: Electron charge misidentification rates in $p_{\text{T}}$ and $\eta$ bins.

| | $\lvert\eta\rvert \leq 1.479$ (B) | $1.479 < \lvert\eta\rvert < 2.5$ (E) |
|---|---|---|
| $10 < p_{\text{T}} < 25$ GeV (L) | $(1.34 \pm 0.41) \times 10^{-4}$ | $(1.99 \pm 0.70) \times 10^{-4}$ |
| $25 < p_{\text{T}} < 50$ GeV (M) | $(2.24 \pm 0.41) \times 10^{-4}$ | $(5.60 \pm 0.41) \times 10^{-4}$ |
| $p_{\text{T}} > 50$ GeV (H) | $(2.28 \pm 0.53) \times 10^{-4}$ | $(13.87 \pm 0.87) \times 10^{-4}$ |

[2]It is assumed that charge identifications are independent for the two electrons. It is also assumed that the electron charge misidentification rate is small so that the probability of misidentifying both electron charges are negligible.

APPENDIX B

# POTENTIAL IMPROVEMENTS IN EVENT CATEGORIES WITH HADRONIC TAUS

In event categories with $\tau_h$'s, the discrimination between the $t\bar{t}H$ signal and major background processes can potentially be improved by adding the helicity information of tau leptons. The spin information of the $\tau$ lepton is preserved in angular and energy distributions of its decay products, especially when it decays hadronically. In case there are two reconstructed $\tau$ leptons, the spin correlation between them provides information regarding the particle that produces the pair the $\tau$ leptons.

## B.1 Tau decays as spin analyzers

A $\tau$ lepton can decay either leptonically into an electron or a muon, each with a branching ratio of about 17%, or hadronically via a scalar or a vector meson. The main $\tau$ decay modes and their branching ratios are summarized in Table B.1. Only pion final states are listed in the table. A $\tau$ lepton can also decay into kaons, but the branching ratios are much smaller.

In case of $\tau^- \rightarrow \pi^- \nu_\tau$, because the pion is a scalar, the neutrino, which is always left-handed, would preferably be emitted in the opposite direction as the tau spin orientation in order to conserve spin. As a result, the charged pion would be emitted preferably to the same direction as the tau spin in the $\tau^-$ rest frame. For a right-handed $\tau^-$, denoted as $\tau^-_R$, (or left-handed $\tau^+$, denoted as $\tau^-_L$), the boost from its rest frame to the lab frame tends to align with the direction of

Table B.1: Main $\tau^-$ decay modes and branching ratios. Branching ratio values are taken from Ref. [62]. Decay modes of $\tau^+$ are the charge conjugate of the $\tau^-$ modes.

| Decay mode | Branching ratio (%) |
|---|---|
| $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$ | $17.39 \pm 0.04$ |
| $\tau^- \to e^- \bar{\nu}_e \nu_\tau$ | $17.82 \pm 0.04$ |
| $\tau^- \to \pi^- \nu_\tau$ | $10.82 \pm 0.05$ |
| $\tau^- (\to \rho^- \nu_\tau) \to \pi^- \pi^0 \nu_\tau$ | $25.49 \pm 0.09$ |
| $\tau^- (\to a_1^- \nu_\tau) \to \pi^- 2\pi^0 \nu_\tau$ | $9.26 \pm 0.10$ |
| $\tau^- (\to a_1^- \nu_\tau) \to \pi^- \pi^+ \pi^- \nu_\tau$ | $8.99 \pm 0.05$ |

its spin, or equivalently the pion momentum, leading to a more energetic pion. On the contrary, a $\tau_L^-$ (or $\tau_R^+$) would preferably decay into a softer pion. In this decay mode, the pion energy, normalized by the energy of $\tau$ lepton, is therefore a good spin indicator.

In case a $\tau$ lepton decays via a spin-1 vector meson ($\rho^-$ or $a_1^-$), the vector meson can be either transversely or longitudinal polarized, corresponding to a helicity state of -1 or 0, respectively. A $\tau_R^-$ (or $\tau_L^+$) is more likely to decay to an energetic longitudinally polarized vector meson, while a $\tau_L^-$ (or $\tau_R^+$) preferably decays to an energetic transversely polarized vector meson. For a transversely polarized vector meson $\rho^-$ or $a_1^-$, its decay products prefer to decay in the collinear direction as its direction of flight, and the energy is more likely to be shared evenly among all decay products. Decay products of a longitudinally polarized vector meson, however, prefer to decay in the direction perpendicular to the direction of flight, and it is more likely one of the decay products carries most of the mother meson energy. Therefore, for final states containing both charged and neutral pions, an observable $\Upsilon$ that quantifies the energy asymmetry between the charged and neutral pions can be used as a spin indicator:

$$\Upsilon = \frac{E_{\pm} - E_0}{E_{\pm} + E_0}, \tag{B.1}$$

where $E_{\pm}$ is the energy of the charged pion and $E_0$ is the energy of the neutral pion(s). Figure B.1 shows the distributions of $\Upsilon$ at generator level for left-handed and right-handed hadronically decaying taus.

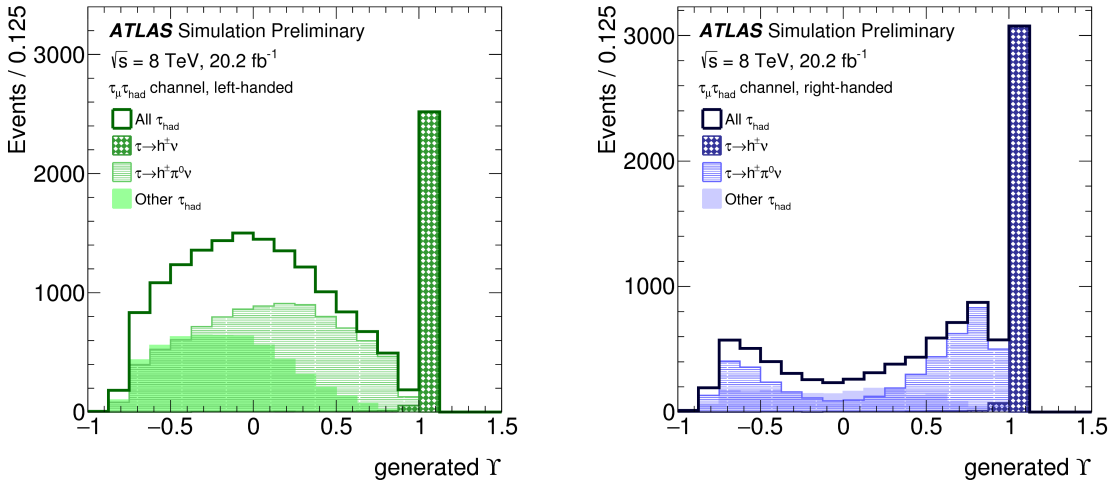


Figure B.1: Distributions of $\Upsilon$ at generator level for $\tau_L^-$ and $\tau_R^+$ (left) and for $\tau_R^-$ and $\tau_L^+$ (right). Plots are taken from Ref. [105].

In case of a final state with three charged pions, in principle angles between each pair of them can be calculated and should carry spin information of the mother $\tau$ lepton. However, a $\tau$ produced from a heavy boson, either the Higgs of Z, is usually highly boosted. As a result, the three charged pion tracks tend to be collimated, which makes it very challenging to measure the angular separations among them given the angular resolution of the tracker. A potentially better observable could be constructed based on the energy asymmetry between the positively and negatively charged pions.

In case of the leptonic $\tau$ decay, spin information is barely accessible due to the production of two neutrinos and also because the helicity of the final state lepton

(e or $\mu$) is not measured in the CMS detector. A more detailed discussion and rigorous calculation regarding the polarization in both hadronic and leptonic tau decays can be found in Ref. [106].

## B.2 Spin correlations in signal and background processes

Because the Higgs boson is a scalar and the Z boson is a vector boson of spin one, the allowed helicity states of the two tau leptons are different in the Higgs and Z decays. In case of $H \to \tau^+\tau^-$, the two taus have to be of the same helicity in order to converse spin. The allowed helicity states from the Higgs boson decay are $\tau_L^+\tau_L^-$ or $\tau_R^+\tau_R^-$. In case of $Z \to \tau^+\tau^-$, the two taus are required to have opposite helicity states, i.e. $\tau_L^+\tau_R^-$ or $\tau_R^+\tau_L^-$. For the background due to misidentified taus, the helicity states of the two tau candidates are expected to be uncorrelated. The spin correlations in different processes can be seen in 2D distributions of the helicity sensitive observables for $\tau^+$ and $\tau^-$.

Figure B.2 shows the energy distribution of the "visible" component (i.e. non-neutrino) of $\tau^+$ and $\tau^-$ from the Higgs in simulated $t\bar{t}H$ events and from the Z boson decay in simulated $t\bar{t}Z$ events. Energies are measured in the mother boson rest frame and are normalized by half of the boson mass. Only the $\tau^\pm \to \pi^\pm\nu_\tau(\bar{\nu}_\tau)$ decay mode is included in the plots.

In case of the decay mode with one charged and one neutral pion in the final state for both $\tau^+$ and $\tau^-$, the 2D distributions of $\Upsilon$ are shown in Fig. B.3 for simulated $t\bar{t}H$ and $t\bar{t}Z$ events. The observable is computed with objects at generator level in the boson rest frame.

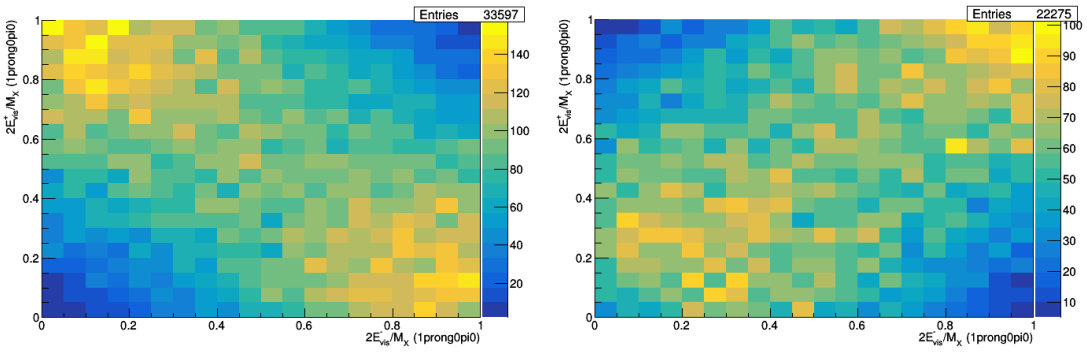


Figure B.2: Distributions of generator level $\tau^+$ and $\tau^-$ visible energies in the mother boson rest frame, normalized by half of the boson mass, for $\tau$ pairs in simulated $t\bar{t}H$ events (left) and $t\bar{t}Z$ events (right). Only the $\tau^\pm \to \pi^\pm \nu_\tau(\bar{\nu}_\tau)$ decay mode is included in the plots.

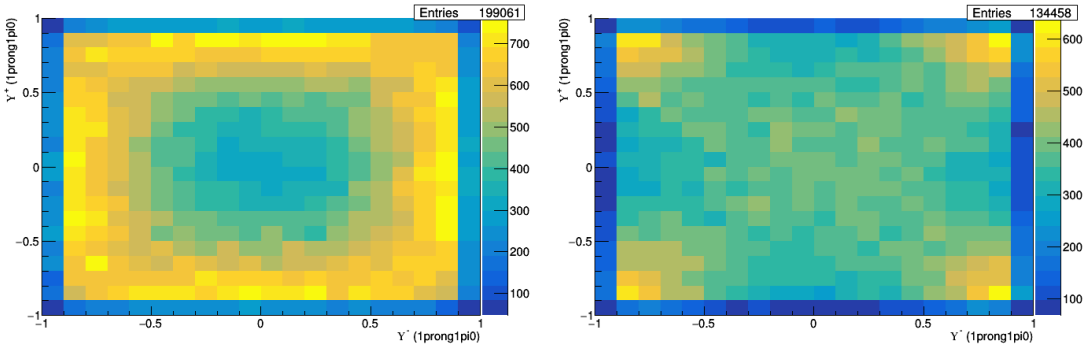


Figure B.3: Distributions of the energy asymmetry of $\tau^+$ and $\tau^-$ in simulated $t\bar{t}H$ (left) and $t\bar{t}Z$ (right) events in the boson rest frame. One the $\tau^\pm \to \pi^\pm \pi^0 \nu_\tau(\bar{\nu}_\tau)$ decay mode in included in the plots.

## B.3 Application and challenges

For the decay mode $\tau^- \to \pi^- \nu_\tau$ and its charge conjugate, the pion energy could be a good spin indicator if the $\tau$ total energy is known or the boson rest frame can be reconstructed. In case of $t\bar{t}H$ events, however, it is very challenging to reconstruct the boson rest frame given its complex final states as well as the production of several neutrinos. The nice spin correlation feature shown in Fig. B.2 is hardly accessible in the lab frame. Nevertheless, in case of the decay modes containing a least one neutral pion, the energy asymmetry in tau decay prod-

ucts is still useful in the lab frame without knowing the rest frame of the tau or the boson. The 2D distributions of the $\Upsilon$ observable in the lab frame are similar to those shown in Fig. B.3 in the boson rest frame.

To study the effect of the helicity sensitive observables on the signal extraction, the $\Upsilon$ observables for $\tau^-$ and $\tau^+$ are added to the input variables of the event-level BDT used in the $1\ell + 2\tau_h$ category. In order to incorporate other $\tau$ decay modes without a neutral pion, $\Upsilon$ is re-expressed in terms of the reconstructed objects in the lab frame as:

$$\Upsilon = \frac{2E_{\text{ldgtrk}}}{E_{\text{vis}}} - 1, \tag{B.2}$$

where $E_{\text{ldgtrk}}$ is the leading track energy and $E_{\text{vis}}$ is the total energy of the visible tau decay products. The new BDT is trained on the same set of simulated $t\bar{t}H$, $t\bar{t}V$, and $t\bar{t}$+jets samples under the same configuration (number of trees, learning rate, maximum depth, etc.) as the nominal one. The two $\Upsilon$ variables for $\tau^-$ and $\tau^+$ rank among the most useful variables to separate signal and background events, after the invariant mass of the $\tau_h$ pair, their $p_T$'s, and the average angular distance between selected jets. The new BDT brings about 7% improvement in the expected upper limit of the $t\bar{t}H$ signal strength in this event category.

Although the preliminary study suggests the variable sensitive to the tau helicity states and their correlation could potentially improve the sensitivity of this analysis, it is also noticed that the improvement is largely due to better discrimination between the $t\bar{t}H$ signal and the $t\bar{t}$+jets background (i.e. misidentified background), rather than the $t\bar{t}V$ background. The $\Upsilon$ variable, which is essentially the leading track energy fraction, has different distributions in events with genuine $\tau$ leptons and events with misidentified $\tau_h$ candidates. Misidentified $\tau_h$'s tend to have softer leading tracks compared to genuine hadronic taus. As

a result, the leading track energy fraction and the $\Upsilon$ are correlated with tau ID working points. The Fake Factor method adopted in this analysis for estimating misidentified backgrounds, however, uses different tau ID working points in the signal region and in the application region. The different distributions of $\Upsilon$ observables, hence the event-level BDT output, in the AR and SR would potentially lead to biases when extrapolating from the AR to estimated the misidentified background in the SR. Such biases could be mitigated by e.g. correcting the tau fake factors or parameterizing the fake factor also as a function of leading track energy. More studies are needed before the tau helicity information can be used in the future version of the analysis.

## BIBLIOGRAPHY


[1] ATLAS Collaboration, "Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] LHC Higgs Cross Section Working Group. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG, last accessed on 2019-06-24.

[4] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector", doi:10.23731/CYRM-2017-002, arXiv:1610.07922.

[5] ATLAS and CMS Collaborations, "Combined measurement of the Higgs boson mass in pp collisions at $\sqrt{s}$ = 7 and 8 TeVwith the ATLAS and CMS experiments", *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.

[6] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV", *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.

[7] ATLAS Collaboration, "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector", *Eur. Phys. J. C* **75** (2015), no. 10, 476, doi:10.1140/epjc/s10052-015-3685-1,10.1140/epjc/s10052-016-3934-y, arXiv:1506.05669. [Erratum: Eur. Phys. J. C 76(2016),no.3,152].

[8] CMS Collaboration, "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV", *Phys. Rev. D* **92** (2015) 012004, doi:10.1103/PhysRevD.92.012004, arXiv:1411.3441.

[9] ATLAS Collaboration, "Constraints on the off-shell Higgs boson signal strength in the high-mass *ZZ* and *WW* final states with the ATLAS

detector", *Eur. Phys. J. C* **75** (2015), no. 7, 335, doi:10.1140/epjc/s10052-015-3542-2, arXiv:1503.01060.

[10] ATLAS Collaboration, "Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb$^{-1}$ of $pp$ collision data", *Phys. Rev. D* **90** (2014), no. 5, 052004, doi:10.1103/PhysRevD.90.052004, arXiv:1406.3827.

[11] CMS Collaboration, "Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs", *Phys. Lett. B* **736** (2014) 64–85, doi:10.1016/j.physletb.2014.06.077, arXiv:1405.3455.

[12] CMS Collaboration, "Limits on the Higgs boson lifetime and width from its decay to four charged leptons", *Phys. Rev. D* **92** (2015), no. 7, 072010, doi:10.1103/PhysRevD.92.072010, arXiv:1507.06656.

[13] ATLAS Collaboration, "Measurements of the Total and Differential Higgs Boson Production Cross Sections Combining the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ Decay Channels at $\sqrt{s}$ = 8 TeV with the ATLAS Detector.", *Phys. Rev. Lett.* **115** (2015), no. 9, 091801, doi:10.1103/PhysRevLett.115.091801, arXiv:1504.05833.

[14] CMS Collaboration, "Measurement of differential and integrated fiducial cross sections for Higgs boson production in the four-lepton decay channel in pp collisions at $\sqrt{s}$ = 7 and 8 TeV", *JHEP* **04** (2016) 005, doi:10.1007/JHEP04(2016)005, arXiv:1512.08377.

[15] CMS Collaboration, "Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s}$ = 8 TeV", *Eur. Phys. J. C* **76** (2016), no. 1, 13, doi:10.1140/epjc/s10052-015-3853-3, arXiv:1508.07819.

[16] ATLAS Collaboration, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector", *Phys. Lett. B* **784** (2018) 173–191, doi:10.1016/j.physletb.2018.07.035, arXiv:1806.00425.

[17] CMS Collaboration, "Observation of $t\bar{t}H$ production", *Phys. Rev. Lett.* **120** (2018), no. 23, 231801, doi:10.1103/PhysRevLett.120.231801, arXiv:1804.02610.

[18] ATLAS Collaboration, "Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in

$pp$ collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector", *Phys. Rev. D* **97** (2018), no. 7, 072016, doi:10.1103/PhysRevD.97.072016, arXiv:1712.08895.

[19] CMS Collaboration, "Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **03** (2019) 026, doi:10.1007/JHEP03(2019)026, arXiv:1804.03682.

[20] ATLAS Collaboration, "Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector", *JHEP* **05** (2016) 160, doi:10.1007/JHEP05(2016)160, arXiv:1604.03812.

[21] CMS Collaboration, "Search for $t\bar{t}H$ production in the all-jet final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **06** (2018) 101, doi:10.1007/JHEP06(2018)101, arXiv:1803.06986.

[22] ATLAS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel using 80 fb$^{-1}$ of $pp$ collision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector", ATLAS CONF Note ATLAS-CONF-2018-028, CERN, Geneva, Jul, 2018.

[23] CMS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **11** (2018) 185, doi:10.1007/JHEP11(2018)185, arXiv:1804.02716.

[24] CMS Collaboration, "An event candidate for the production of a top quark and top anti-quark pair in conjunction with a Higgs Boson in the CMS detector.", (Jun, 2018). CMS-PHO-EVENTS-2018-004.

[25] ATLAS Collaboration, "Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector", *Phys. Rev. D* **97** (2018), no. 7, 072003, doi:10.1103/PhysRevD.97.072003, arXiv:1712.08891.

[26] CMS Collaboration, "Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying $\tau$ leptons at $\sqrt{s}$ = 13 TeV", *JHEP* **08** (2018) 066, doi:10.1007/JHEP08(2018)066, arXiv:1803.05485.

[27] P. Langacker, "The standard model and beyond". 2010. ISBN 9781420079067,

[28] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.

[29] F. Marcastel, "CERN's Accelerator Complex. La chaîne des accélérateurs du CERN",. OPEN-PHO-CHART-2013-001.

[30] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.

[31] ATLAS Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[32] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[33] LHCb Collaboration, "The LHCb detector at the LHC", *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

[34] CMS Collaboration, "The CMS tracker: addendum to the Technical Design Report", Technical Report CERN-LHCC-2000-016, 2000.

[35] CMS Collaboration, "CMS Technical Design Report for the Pixel Detector Upgrade", Technical Design Report CERN-LHCC-2012-016, CMS-TDR-011, FERMILAB-DESIGN-2012-02, Sep, 2012. doi:10.2172/1151650.

[36] CMS Collaboration, "The CMS electromagnetic calorimeter project: Technical Design Report", Technical Design Report CERN-LHCC-97-033, CMS-TDR-4, 1997.

[37] CMS Collaboration, "The CMS hadron calorimeter project: Technical Design Report", Technical Design Report CERN-LHCC-97-031, CMS-TDR-2, 1997.

[38] CMS Collaboration, "CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter", Technical Design Report CERN-LHCC-2012-015, CMS-TDR-010, 2012. doi:10.2172/1151651.

[39] CMS Collaboration, "The CMS muon project: Technical Design Report", Technical Design Report CERN-LHCC-97-032, CMS-TDR-3, 1997.

[40] CMS Collaboration, "The Performance of the CMS Muon Detector in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV at the LHC", *JINST* **8** (2013) P11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.

[41] CMS Collaboration, "CMS The TriDAS project: Technical design report, vol. 1: The trigger systems", Technical Design Report CERN-LHCC-2000-038, CMS-TDR-6-1, 2000.

[42] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017), no. 01, P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[43] CMS Collaboration, "CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project", Technical Design Report CERN-LHCC-2002-026, CMS-TDR-6, 2002.

[44] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[45] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[46] CMS Collaboration, "2017 tracking performance plots", CMS Performance Note CMS-DP-2017-015, CERN, April, 2017.

[47] P. Billoir, "Progressive track recognition with a Kalman-like fitting procedure", *Comput. Phys. Commun.* **57** (1989), no. 1, 390 – 394, doi:https://doi.org/10.1016/0010-4655(89)90249-X.

[48] P. Billoir and S. Qian, "Simultaneous pattern recognition and track fitting by the Kalman filtering method", *Nucl. Instrum. Meth. A* **294** (1990), no. 1, 219 – 228, doi:https://doi.org/10.1016/0168-9002(90)91835-Y.

[49] R. Mankel, "A Concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system", *Nucl. Instrum. Meth. A* **395** (1997) 169–184, doi:10.1016/S0168-9002(97)00705-5.

[50] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth. A* **262** (1987) 444–450, doi:10.1016/0168-9002(87)90887-4.

[51] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-$k_t$ jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[52] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P0601, arXiv:1804.04528.

[53] CMS Collaboration, "Muon identification and isolation efficiencies with 2017 and 2018 data", CMS Performance Note CMS-DP-2018-042, Jul, 2018.

[54] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s}$ = 8 TeV", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[55] H. Bethe, W. Heitler, and P. A. M. Dirac, "On the stopping of fast particles and on the creation of positive electrons", *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **146** (1934), no. 856, 83–112, doi:10.1098/rspa.1934.0140.

[56] CMS Collaboration, "Electron and Photon performance in CMS with the full 2017 data sample and additional 2016 highlights for the CALOR 2018 Conference", CMS Performance Note CMS-DP-2018-017, May, 2018.

[57] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[58] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[59] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2017.

[60] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018), no. 05, P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[61] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[62] Particle Data Group Collaboration, "Review of particle physics", *Phys. Rev. D* **98** (Aug, 2018) 030001, doi:10.1103/PhysRevD.98.030001.

[63] CMS Collaboration, "Performance of reconstruction and identification of $\tau$ leptons decaying to hadrons and $\nu_\tau$ in pp collisions at $\sqrt{s}$ = 13 TeV", *JINST* **13** (2018), no. 10, P10005, doi:10.1088/1748-0221/13/10/P10005, arXiv:1809.02816.

[64] CMS Collaboration, "Reconstruction and identification of tau lepton decays to hadrons and tau neutrino at CMS", *JINST* **11** (2016), no. 01, P01019, doi:10.1088/1748-0221/11/01/P01019, arXiv:1510.07488.

[65] CMS Collaboration, "Performance of missing transverse momentum in pp collisions at sqrt(s)=13 TeV using the CMS detector", CMS Physics Analysis Summary CMS-PAS-JME-17-001, 2018.

[66] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[67] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[68] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[69] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations", *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.

[70] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[71] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEGmethod", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[72] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[73] R. Frederix, E. Re, and P. Torrielli, "Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO", *JHEP* **09** (2012) 130, doi:10.1007/JHEP09(2012)130, arXiv:1207.5391.

[74] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[75] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, "$W^+W^-$, WZ and ZZ production in the POWHEG BOX", *JHEP* **11** (2011) 078, doi:10.1007/JHEP11(2011)078, arXiv:1107.5051.

[76] S. Frixione, P. Nason, and G. Ridolfi, "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[77] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[78] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A brief introduction to PYTHIA8.1", *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[79] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", arXiv:1903.12179.

[80] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[81] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119–126, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[82] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.

[83] CMS Collaboration, "Measurements of Inclusive $W$ and $Z$ Cross Sections in $pp$ Collisions at $\sqrt{s}$ = 7 TeV", *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.

[84] K. Kondo, "Dynamical Likelihood Method for Reconstruction of Events With Missing Momentum. 1: Method and Toy Models", *J. Phys. Soc. Jap.* **57** (1988) 4126–4140, doi:10.1143/JPSJ.57.4126.

[85] K. Kondo, "Dynamical likelihood method for reconstruction of events with missing momentum. ii. mass spectra for 2 to 2 processes", *Journal of the Physical Society of Japan* **60** (1991), no. 3, 836–844, doi:10.1143/JPSJ.60.836, arXiv:https://doi.org/10.1143/JPSJ.60.836.

[86] R. H. Dalitz and G. R. Goldstein, "Decay and polarization properties of the top quark", *Phys. Rev. D* **45** (Mar, 1992) 1531–1543, doi:10.1103/PhysRevD.45.1531.

[87] G. R. Goldstein, K. Sliwa, and R. H. Dalitz, "Observing top-quark production at the fermilab tevatron", *Phys. Rev. D* **47** (Feb, 1993) 967–972, doi:10.1103/PhysRevD.47.967.

[88] ATLAS Collaboration, CMS Collaboration, The LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011.

[89] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra", in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, pp. 115–120. 2011. arXiv:1103.0354. doi:10.5170/CERN-2011-006.115.

[90] CMS Collaboration, "Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **08** (2018) 011, doi:10.1007/JHEP08(2018)011, arXiv:1711.02547.

[91] ATLAS Collaboration, "Measurement of the $W^{\pm}Z$ boson pair-production cross section in $pp$ collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector", *Phys. Lett. B* **762** (2016) 1–22, doi:10.1016/j.physletb.2016.08.052, arXiv:1606.04017.

[92] CMS Collaboration, "Measurement of the WZ production cross section in pp collisions at $\sqrt{(s)}$ = 13 TeV", *Phys. Lett. B* **766** (2017) 268–290, doi:10.1016/j.physletb.2017.01.011, arXiv:1607.06943.

[93] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s}$ = 13 TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, CERN, Geneva, 2018.

[94] G. Apollinari et al., "High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1". CERN Yellow Reports: Monographs. CERN, Geneva, 2017. doi:10.23731/CYRM-2017-004.

[95] CMS Collaboration, "Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid", CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, CERN, 2015.

[96] HL-LHC project team. http://hilumilhc.web.cern.ch/about/hl-lhc-project, last accessed on 2019-07-03.

[97] C. Foudas, A. Rose, J. Jones, and G. Hall, "A Study for a tracking trigger at first level for CMS at SLHC", in *Proceedings, eleventh Workshop on Electronics for LHC and Future Experiments, Heidelberg, Germany, 12-16 September 2005*, p. 90. 2005. arXiv:physics/0510227.

[98] CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker",

Technical Design Report CERN-LHCC-2017-009. CMS-TDR-014, CERN, Geneva, Jun, 2017.

[99] E. Bartz et al., "FPGA-Based Tracklet Approach to Level-1 Track Finding at CMS for the HL-LHC", *EPJ Web Conf.* **150** (2017) 00016, doi:10.1051/epjconf/201715000016, arXiv:1706.09225.

[100] R. Aggleton et al., "An FPGA based track finder for the L1 trigger of the CMS experiment at the High Luminosity LHC", *JINST* **12** (2017), no. 12, P12019, doi:10.1088/1748-0221/12/12/P12019.

[101] G. Fedi, "Associative memory pattern matching for the L1 track trigger of CMS at the HL-LHC", *Eur. Phys. J. Conf.* **127** (2016) 7, doi:10.1051/epjconf/201612700008.

[102] A. Svetek et al., "The Calorimeter Trigger Processor Card", *JINST* **11** (2016) C02011, doi:10.1088/1748-0221/11/02/C02011.

[103] Xilinx Incorporated, "7 Series FPGAs Overview". https://www.xilinx.com/products/silicon-devices/fpga/virtex-7.html, last accessed on 2017-03-28.

[104] Boston University, "The AMC13 Project". http://bucms.bu.edu/twiki/bin/view/BUCMSPublic/HcalDTC, last accessed on 2017-03-28.

[105] ATLAS Collaboration, "Measurement of Tau Polarisation in $Z/\gamma^* \to \tau\tau$ Decays in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector", ATLAS Note ATLAS-CONF-2017-049, CERN, Geneva, Jul, 2017.

[106] B. Bullock, K. Hagiwara, and A. Martin, "Tau polarization and its correlations as a probe of new physics", *Nuclear Physics B* **395** (1993), no. 3, 499 – 533, doi:https://doi.org/10.1016/0550-3213(93)90045-Q.